

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷: A61K 48/00	A2	(11) International Publication Number: WO 00/25827 (43) International Publication Date: 11 May 2000 (11.05.00)
(21) International Application Number: PCT/EP99/07874 (22) International Filing Date: 18 October 1999 (18.10.99) (30) Priority Data: MI98A002330 30 October 1998 (30.10.98) IT (71) Applicant (for all designated States except US): MENARINI RICERCHE S.P.A. [IT/IT]; Via Tito Spert, 10, I-00040 Pomezia (IT). (72) Inventors; and (75) Inventors/Applicants (for US only): PARIENTE, Dino [IT/IT]; (IT), DI MASSIMO, Anna, Maria [IT/IT]; (IT), DE SANTIS, Rita [IT/IT]; Via Rismondo, I-50131 Firenze (IT). (74) Agent: MINOIA, Fabrizio; Bianchetti Bracco Minoia Srl, Via Rossini, 8, I-20122 Milano (IT).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GR, GM, GU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: PHARMACEUTICAL COMPOSITION, CONTAINING FRAGMENTS OF AN ANTIGENIC PROTEIN ENCODING DNA ENDOWED WITH ANTI-TUMOR EFFECT (57) Abstract Provided herein is a pharmaceutical composition containing one or more DNA molecules encoding fragments of a protein overexpressed in tumor cells, in order to induce an anti-tumor Ag-specific immune response, in association with suitable excipients and adjuvants.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SE	Sweden
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	SD	Sudan
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TD	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KP	Democratic People's Republic of Korea	NZ	New Zealand	ZW	Zimbabwe
CI	Cote d'Ivoire	KR	Republic of Korea	PL	Poland		
CM	Cameroon	KZ	Kazakhstan	PT	Portugal		
CN	China	LC	Saint Lucia	RO	Romania		
CU	Cuba	LI	Liechtenstein	RU	Russian Federation		
CZ	Czech Republic	LK	Sri Lanka	SD	Sudan		
DE	Germany	LR	Liberia	SE	Sweden		
DK	Denmark			SG	Singapore		
EE	Estonia						

PHARMACEUTICAL COMPOSITION, CONTAINING FRAGMENTS OF AN ANTIGENIC PROTEIN ENCODING DNA ENDOWED WITH ANTI-TUMOR EFFECT.

Field of the invention

5 The invention relates to a pool of DNA plasmid constructs containing the sequences of human MUC-1 encoding fragments and to a pool of DNA plasmids in which the fragments themselves are preceded by the sequence encoding a protein consisting of human ubiquitin fused to a bacterial LacI fragment. The invention
10 further relates to their use in the preparation of pharmaceutical compositions for use as DNA anti-tumor vaccines.

Background art

The invention provides an anti-tumor therapy based on the induction or activation of the immune response able to bring
15 about tumor rejection. The validity of such an idea is demonstrated from the first clinical results; for example, patients treated with a viral vaccine containing the Carcinoembryonic Antigen (CEA) encoding sequences demonstrated immune system activation against this antigen (Tsang KY et al.
20 J. Natl. Cancer. Inst. 87: 982, 1995).

The activation of an immune anti-tumor response is achievable through four different approaches:

- a) Ex vivo engineering of patient tumor cells in order to make them more immunogenic and suitable as a vaccine;
- 25 b) Ex vivo engineering of patient immune cells in order to pre-activate an *in vitro* immune response.
- c) Inoculation of naked or liposome capsulated or viral particle integrated (retrovirus, vaccinia virus, adenovirus, etc.) DNA encoding tumor associated antigens;
- 30 d) Treatment with recombinant or synthetic soluble tumor antigens conjugated or mixed with adjuvants.

The first two approaches consist of the engineering of every single patient cell and are limited in that they are necessarily patient-specific, while the latter two are aimed to

obtain products comparable to a traditional drug.

The new vaccination methods reflect the development of new technologies. The recent indications coming from the experimentation on DNA naked vaccines that induce either a persistent antibody or a cell immune response, make the traditional protein subunit vaccines constituted of certain specific peptides, inducing a lymphocyte population, obsolete. Intramuscularly or intradermically injected proteins, encoded by naked DNA, induce a cytotoxic-specific response as well as a helper response. This powerful combination is extremely effective but the underling mechanism is not completely clarified yet. Muscle cells express class I MHC antigens at low levels only, and do not apparently express class II antigens or co-stimulatory molecules. Consequently, transfected muscle cells are unlikely to play an important role in the onset of the immune response per se. Recent data show that Antigen Presenting Cells (APC), such as macrophages or dendritic cells, play a fundamental role in capturing the myocyte released antigen and in the subsequent processing and presenting of the respective peptides in the context of the class I and II molecules, thus inducing a CD8+ cell activation with cytotoxic activity as well as activation of the CD4+ cells co-operating with B lymphocytes in eliciting the antibody response (Corr M et al *J. Exp. Med.* 184:1555, 1996) (Tighe, H. et al. *Immunology Today* 19:89, 1998).

Furthermore, the use of cytokines is known to improve the therapeutic effect deriving from immunization with DNA. Cytokines can be administered in the form of exogenous proteins as reported in Irvine et al., *J. Immunol.* 156: 238, 1996. An alternative approach is represented by the contemporaneous inoculation of both the tumor antigen or the desired cytokine encoding plasmids, thus allowing the cytokine to be produced in situ (Kim JJ et al. *Immunol* 158: 816, 1997).

The active immunization approach of the present invention is based on the use of DNA vectors as vaccines against the MUC-1

human antigen or Polymorphic Epithelial Mucin (PEM), overexpressed in tumor cells. MUC-1 is an epithelial luminal surface glycoprotein (Patton S. et al. *BBA* 1241:407, 1995). In the cell transformation process this glycoprotein loses the apical localization and its expression level rises dramatically. The protein function consists of protecting the luminal surfaces, for example in the mammary gland, ovary, endometrium, colon, stomach, pancreas, bladder, kidney, etc. A glycosylation defect is reported that makes tumor cell associated MUC-1 antigenically different from normal cell associated MUC-1. This phenomenon causes tumor MUC-1 to expose the antigen epitopes that are normally masked by the sugar moieties in the normal cell expressed MUC-1. This characteristic makes tumor MUC-1 particularly interesting in an induction of a tumor specific antibody response (Apostolopoulos V. et al. *Crit. Rev. Immunol.* 14:293, 1994).

As an objective, the vaccination is aimed at inducing immune responses against tumor cells expressing MUC1 at high levels, preserving at the same time the low expressing normal epithelia. The DNA vaccination relies upon the entrance of a gene or portions thereof inside the body cells followed by transcription and translation of the inserted sequence and thus the intracellular synthesis of the corresponding polypeptide. An important advantage of this system is that the neo-synthesized protein is naturally processed inside the cell and the produced peptides are associated with the Major Histocompatibility Complex class I molecules (MHC-I). The MHC/peptide complexes are therefore naturally exported to the cell surface where they can be recognized by the immune system CD8+ cytotoxic cells. Only the polypeptides synthesized inside the cell are then processed and presented in association with the MHC class I molecules, thus making it the only mechanism to stimulate, a specific cytotoxic response. Vaccination systems based on protein or peptide administration are usually more effective in stimulating

the antibody immune response which, to date, has been shown to be ineffective in rejecting tumor cells. Current gene therapy techniques rely upon DNA packaging in recombinant viral vectors (retrovirus and adenovirus). The naked DNA administration is much more advantageous in terms of effectiveness and safety compared to viral vector therapies (Kumar V and Sercarz E. *Nature Med.* 2: 857, 1996; McDonnell WM et al., *New England J. of Med.* 334: 42, 1996). In fact naked DNA is unable either to duplicate or integrate in the host tissue DNA and does not induce the immune response to viral proteins.

The use of the ubiquitin to enhance the neo-synthesized protein processing and thus cytotoxic lymphocyte induction was recently reported (Rodriguez F. et al., *J. Virology* 71: 8497, 1997). The use of ubiquitin in order to generate proteins with an N-terminal amino acid, making them unstable and thus prone to enhanced degradation, had been previously reported (Bechmair A. et al., *SCIENCE* 234: 179, 1986). The higher instability of these proteins was subsequently related to enhanced intracellular processing and presentation of model proteins by MHC-1 (Grant E P et al., *J. Immunol.* 155: 3750, 1995) (Wu Y and Kipps T.J., *J. Immunol.* 159: 6037, 1997).

The use of single constructs containing partial antigen encoding DNA fragments (influenza virus nucleoprotein), having a higher antigenic presentation efficiency compared to the analogues with the whole antigenic sequence, in DNA vaccination was reported (Anton L. C. et al., *J. Immunol.* 158: 2535, 1997). Furthermore the processing of intracellular proteins and presentation of the respective peptides by MHC class I proteins in physiologic conditions, underlie the mechanism of immunological surveillance. For a given protein and a specific MHC context, there are peptide fragments termed dominants (i. e. prevailing on subdominants or cryptics), which are unable to generate any immune response because they are recognized as "self". It has now been outlined, according to an aspect of the

present invention, that an approach aimed at supporting the non-dominant epitope presentation by the administration of a mix of antigen protein fragments is able to elicit a surprising cytotoxic immune response.

5 Description of the invention

It has now been found that DNA molecules, encoding fragments of a protein overexpressed in tumor cells, can be conveniently used to induce an antigen-specific anti-tumor immune response.

10 The invention relates particularly to a pharmaceutical composition containing one or more DNA encoding Mucin (MUC-1) protein fragments.

 The DNA used in the present invention can be plasmid or viral DNA, preferably plasmid DNA obtained employing the pMRS30
15 expression vector described in fig. 13.

 The compositions according to the invention contain preferably at least two DNA fragments of the Mucin (MUC-1) or of another protein overexpressed in tumor cells.

 The compositions according to the invention contain
20 preferably at least four fragments, each ranging from 200 to about 700 nucleotides, each sequence being juxtaposed and possibly partially overlapping, from about 50 to about 150 nucleotides, at the 3' and/or 5' end of the adjacent one.

 The DNA fragments according to the invention can be
25 possibly preceded at the 5' end by a ubiquitin encoding DNA sequence and possibly also by a LacI portion of Escherichia coli.

 The invention relates also to new DNA fragments and to the use of Mucin-1 fragments defined above in the medicine and anti-
30 tumor vaccine preparation.

Description of the figures

 Fig. 1

 Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS166 expression

vector. This DNA includes the sequence corresponding to nucleotides 136-339 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by the two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the amino acids encoded by the 136-339 fragment of the EMBL sequence J05581.

Fig. 2

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS169 expression vector. This DNA includes the sequence corresponding to nucleotides 205-720 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the amino acids encoded by the 205-720 fragment of the EMBL sequence J05581.

Fig. 3

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS168 expression vector. This DNA includes the sequence corresponding to nucleotides 631-1275 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the amino acids encoded by the 631-1275 fragment of the EMBL sequence J05581.

Fig. 4

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS167 expression vector. This DNA includes the sequence corresponding to nucleotides 1222-1497 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the

amino acids encoded by the 1222-1497 fragment of the EMBL sequence J05581.

Fig. 5

5 Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS175 expression vector. This DNA includes the sequence corresponding to nucleotides 136-1497 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The
10 encoded polypeptide thus includes a Metionin followed by the amino acids encoded by the 136-1497 fragment of the EMBL sequence J05581.

Fig. 6

15 Nucleotide DNA sequence (with the respective amino acid sequence) termed UBILacI. The encoded polypeptide includes the Ubiquitin sequence fused to a partial sequence of the bacterial protein beta-galactosidase, as described in Chau V. et al. *Science* 243: 1576, 1989.

Fig. 7

20 Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the expression vector pMRS30 to give the pMRS171 expression vector. This DNA includes the sequence termed UBILacI (see fig. 6) fused to the sequence corresponding to nucleotides 136-339 of the EMBL sequence J05581
25 followed by two translation stop codons, TGA and TAA. The coded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 136-339 of the EMBL sequence J05581.

Fig. 8

30 Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS174 expression vector. This DNA includes the sequence termed UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 205-720 of the EMBL

sequence J05581 followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 205-720 of the EMBL sequence J05581.

Fig. 9

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS173 expression vector. This DNA includes the sequence termed UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 631-1275 of the EMBL sequence J05581 followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 631-1275 of the EMBL sequence J05581.

Fig. 10

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS172 expression vector. This DNA includes the sequence termed UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 1222-1497 of the EMBL sequence J05581 followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 1222-1497 of the EMBL sequence J05581.

Fig. 11

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS176 expression vector. This DNA includes the sequence named UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 136-1497 of the EMBL sequence J05581 followed by two translation stop codons, TGA and

TAA. The encoded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 136-1497 of the EMBL sequence J05581.

5 **Fig. 12**

Electrophoretic analysis on 1% agarose gel in 1X TBE. mRNA extracted from CHO, CD34+ dendritic cells and dendritic cells from PBMC, respectively, transfected with pMRS169, and subjected to RT-PCR reaction either with (lanes 4, 8, 12) or without (lanes 5, 9, 13) Reverse Transcriptase. Molecular weight DNA marker (lane 1); internal negative controls (lanes 2, 6); internal positive controls (lanes 3, 7, 10, 11); positive control from Promega kit (lane 14).

10 **Fig. 13**

Nucleotide sequence of the pMRS30 expression vector. The 1-2862 region corresponds to the AccI (location 504) - BamHI (location 3369) region of the pSV2CAT vector (EMBL M77788); the 2863-3721 region includes the human cytomegalovirus promoter (human cytomegalovirus major immediate-early gene enhancer); the 3722-4905 region includes several cloning sites, including XbaI (location 3727), and the processing signal of the rabbit beta-globin gene.

15 **Detailed description of the invention**

A DNA plasmid pool encoding, in eukaryotic cells, fragments of the MUC-1 human protein antigen was prepared. Constructs are based on the mammalian expression vector termed pMRS30, described in figure 13 and previously claimed in the Patent Application W095/11982, and contain partial sequences of the MUC-1 cDNAs reported in the EMBL database with accession number J05581. MUC-1 encoding DNA was fragmented so that each fragment represents a discrete portion, partially overlapping to the adjacent ones. Administration of a mix of such plasmids can cause different plasmids to transfect different APC cells at the administration site. Therefore such cells produce and process

discrete portions of the MUC-1 protein giving the related peptides. In those conditions, the occurring subdominant and cryptic peptides can also be presented in association with class I MHC molecules thus generating a cytotoxic immune response.

5 The present invention thus relates to the use of a group of four constructs (Figures 1 to 4) containing MUC-1 cDNA partial fragments in admixture containing at least two of them and a group of four constructs (Figures 7 to 10) containing MUC-1 cDNA partial fragment preceded by the DNA encoding a protein sequence
10 containing Ubiquitin and an Escherichia coli Lac I portion (Figure 6) used separately or in admixture containing at least two of them.

The present invention relates also to the use of the construct (Figure 5) containing the almost complete sequence of
15 the MUC-1 cDNA and the construct (Figure 11) containing the almost complete sequence of the MUC-1 cDNA preceded by the DNA encoding a protein sequence containing Ubiquitin and an Escherichia coli Lac I portion.

20 The mixture of the four constructs containing the partial fragments of the MUC-1 cDNA and the mixture of the four constructs containing the partial fragments of the MUC-1 cDNA preceded by the DNA encoding a protein sequence, containing Ubiquitin and an Escherichia coli Lac I portion, represents a preferred embodiment of the present invention.

25 Constructs according to the present invention can be used in the anti-tumor therapy of patient affected with tumors characterized by high MUC-1 expression.

Constructs described in the present invention were obtained as follows.

30 In the case of the first series of constructs, the fragments of the MUC-1 DNA were obtained by RT-PCR from BT20 cell line or by DNA partial chemical synthesis. Such fragments were then cloned into the pMRS30 expression vector and verified by sequencing.

In the case of the second series of constructs, the fragments were obtained from the first series of constructs by a PCR re-amplification. These fragments were then fused to the DNA encoding the Ubiquitin (obtained by RT-PCR from MCF7 cell line mRNA) and a partial lacI sequence (obtained by PCR from the commercial vector pGEX). DNA sequences thus obtained were then cloned in the pMRS30 expression vector and verified by sequencing. For the intended therapeutic or prophylactic uses, fragments or constructs according to the invention are suitably formulated, using carriers and methods previously employed in naked DNA vaccines, as described for example in The Immunologist, 1994, 2:1; WO 90/11092, Proc. Natl. Acad. Sci. U.S.A., 1986, 83, 9551; US 5580859; Immunology today 19 (1998), 89-97; Proc. Natl. Acad. Sci. U.S.A. 90 (1993), 11478-11482; Nat. Med. 3 (1997), 526-532; Vaccine 12 (1994), 1495-1498; DNA Cell. Biol. 12 (1993), 777-783. The dosages will be determined on the basis of clinical and pharmacological-toxicological trials. Generally speaking, they will be comprised between 0.005 µg/kg and 5 µg/kg of the fragment mix. The composition of the invention can also contain a cytokine or a cytokine encoding plasmid.

The invention will be further illustrated by means of the following examples.

Example 1. Plasmid pMRS166 construction.

BT20 tumor cells (ATCC HTB-19) were cultured in Eagles MEM supplemented with 10% fetal calf serum. Ten million cells were trypsinized, washed with PBS, and mRNA extracted.

An aliquot of this RNA was subjected to RT-PCR (reverse transcriptase-polymerase chain reaction) reaction in the presence of the following synthetic oligonucleotides:

V11 (5' GATCTCTAGAAATGACAGGTTCTGGTCATGCAAGC 3')

V4 (5' GATCTCTAGAAAGCCTTATCAACCTGAAGCTGGTTCGGTGGC 3')

The produced DNA fragment, purified and digested with the restriction enzyme XbaI, was cloned into the pMRS30 expression

vector, containing the human cytomegalovirus promoter and the beta-globin polyadenylation signal as claimed in the Patent WO9511982. The resulting pMRS166 vector contains a DNA fragment including the ATG codon, the sequence corresponding to the nucleotides 136-339 of the EMBL sequence J05581, and two stop codons, TGA and TAA.

This fragment is reported in fig. 1.

Example 2. Plasmid pMRS169 construction.

An aliquot of the RNA obtained as reported in example 1 was amplified by RT-PCR in the presence of the following synthetic oligonucleotides:

V12 (5 GATCTCTAGAATGGTGGCCAGCTCTACTGAGAAGAATGC 3)

V15 (5 GGGCGTGGAGCCCGGCTGGCTTGT 3)

The produced DNA fragment, purified and digested with the restriction enzymes SmaI and XbaI, was fused, by the SmaI restriction site, to a DNA fragment entirely synthetically constructed, and including a sequence partially corresponding to the nucleotides 457-720 of the EMBL sequence J05581 and two stop codons, TGA and TAA. The whole fragment was thus cloned in the XbaI site of the pMRS30 expression vector. The resulting pMRS169 vector contains a DNA fragment including the ATG codon, the sequence partially corresponding to the nucleotides 205-720 of the EMBL sequence J05581, and two stop codons, TGA and TAA.

This fragment is reported in fig. 2.

Example 3. Plasmid pMRS168 construction.

An aliquot of the RNA obtained as reported in example 1 was amplified by RT-PCR in the presence of the following synthetic oligonucleotides:

V13 (5 GATCTCTAGAATGGGCTCAGCTTCTACTCTGGTGACAAACGGC 3)

V8 (5 GATCTCTAGAAAGCTTATCACAAGGCAATGAGATAGACAAATGCC 3)

The produced DNA fragment, purified and digested with the restriction enzyme XbaI was cloned in the pMRS30 expression vector. The resulting pMRS168 vector contains a DNA fragment including the ATG codon, the sequence corresponding to the

nucleotides 631-1275 of the EMBL sequence J05581, and two stop codons, TGA and TAA.

This fragment is reported in fig. 3.

Example 4. Plasmid pMRS167 construction.

- 5 An aliquot of the RNA obtained as reported in example 1 was subjected to RT-PCR reaction in the presence of the following synthetic oligonucleotides:

V14 (5 GATCTCTAGAATGCTGGTCTGGTCTGTGTTCTGTGCC 3)

V10 (5 GATCTCTAGAAAGCTTATCACAAGTTGGCAGAAAGTGGCTGC 3)

- 10 The produced DNA fragment, purified and digested with the restriction enzyme XbaI was cloned in the pMRS30 expression vector. The resulting pMRS167 vector contains a DNA fragment including the ATG codon, the sequence corresponding to the nucleotides 1222-1497 of the EMBL sequence J05581, and two stop
15 codons, TGA and TAA.

This fragment is reported in fig. 4.

Example 5. Plasmid pMRS175 construction.

pMRS166, 169, 168, 167 plasmids were subjected to PCR reaction in the presence of the following nucleotide pairs:

- 20 V11 (see example 1)
V18 (5 AACCTGAAGCTGGTTCOGTGGC 3) for pMRS166
V19 (5 GTGCCAGCTCTACTGAGAGAATGC 3)
V20 (5 GCTGGGAATTGAGAATGGAGTGCTCTTGC 3) for pMRS169
V21 (5 GGCTCAGCTTCTACTCTGGTGACACAAGGCC 3)
25 V22 (5 CAAGGCAATGAGATAGACAATGGCC 3) for pMRS168
V23 (5 CTGGTGCTGGTCTGTGTTCTGGTTGCG 3)
V10 (see example 4) for pMRS167

- The four DNA fragments obtained in the respective PCR reactions were mixed in equimolar amounts and PCR reacted in the presence of the V11 and V10 oligonucleotides.
30

The produced DNA fragment, purified and digested with the XbaI restriction enzyme, was cloned in the pMRS30 expression vector. The resulting pMRS175 vector contains a DNA fragment including the ATG codon, the sequence partially corresponding to

the nucleotides 136-1497 of the EMBL sequence J05581 and two stop codons TGA and TAA.

This fragment is reported in fig. 5.

Example 6. Plasmid pMRS171 construction.

- 5 MCF7 tumor cells (ATCC HTB-22) were cultured in Eagles MEM supplemented with 10% fetal calf serum. Ten million cells were trypsinized, washed with PBS, and mRNA extracted.

An aliquot of this RNA was subjected to RT-PCR in the presence of the following synthetic oligonucleotides:

- 10 UBIup (5GATCTCTAGAATGCAGATCTTCGTGAAGACCTTGACTGGT 3)
UBIdown
(5TCACCAGCGAGAGCGGGCAACAGCCATGCACCACTACCGTCCTCCACCTCTGAGACGGAGC
ACCAAG 3)

The reaction produces a DNA fragment termed fragment 1.

- 15 DNA from pGEX11T (Pharmacia) was subjected to PCR reaction in the presence of the following synthetic oligonucleotides:

LacIup (5CCTCGTCTCAGAGGTGGGAGGCACGGTAGTGGTCATGGCTGTTGCC
GTCTCGCTGGTGAAAAG 3)
LacIdown (5GATCGGATCCTCGGGAAACCTGTCTGTCAGCTGC 3)

- 20 This reaction gives a DNA fragment termed fragment 2.

The 1 and 2 DNA fragments, obtained in the respective PCR reactions, were mixed in equimolar amounts and subjected to PCR reaction in presence of the UBIup and LacIdown oligonucleotides.

- 25 The produced DNA fragment, purified and digested with the restriction enzymes XbaI and BamHI, was cloned into the pUC18 commercial plasmid. The resulting pMRS156 vector contains a DNA fragment including the sequence encoding the ubiquitin fused to the sequence encoding a bacterial beta-galactosidase portion. This fragment, termed UBILacI, is reported in fig. 6.

- 30 Plasmid pMRS166 DNA was subjected to a PCR reaction in presence of the following synthetic oligonucleotides:

V3 (5GATCGGATCCACAGGTTCTGGTCAAGC 3)

V4 (see Example 1)

The produced DNA fragment, purified and digested with the

restriction enzymes XbaI and BamHI, was fused, by ligation into the two BamHI sites, to the UBILacI fragment deriving from the pMRS156 plasmid. The resulting fragment was cloned into the pMRS30 expression vector. The resulting pMRS171 vector contains a DNA fragment including the UBILacI sequence, the sequence corresponding to the 136-339 nucleotides of the EMBL sequence J05581 and two stop codons, TGA and TAA. This fragment is reported in fig. 7.

Example 7. Plasmid pMRS174 construction.

Plasmid pMRS169 DNA was subjected to PCR reaction in the presence of the following synthetic oligonucleotides:

V5 (5GATCGGATCCGTCGCCAGCTCTACTGAGAAGATGC 3)

V6 (5GATCTCTAGAAAGCITATCAGCTGGGAATTGAGAATGGAGTGCTCTTC 3)

The produced DNA fragment, purified and digested with the restriction enzymes XbaI and BamHI, was fused, by ligation into the two BamHI sites, to the UBILacI fragment deriving from the pMRS156 plasmid. The resulting fragment was cloned into the pMRS30 expression vector. The resulting pMRS174 vector contains a DNA fragment including the UBILacI sequence, the sequence corresponding to the 205-720 nucleotides of the EMBL sequence J05581, and two stop codons, TGA and TAA. This fragment is reported in fig. 8.

Example 8. Plasmid pMRS173 construction.

Plasmid pMRS168 DNA was subjected to PCR reaction in the presence of the following synthetic oligonucleotides:

V7 (5GATCGGATCCGGCTCAGCTTCTACTCTGGTGACACACGGC 3)

V8 (see example 3)

The produced DNA fragment, purified and digested with the restriction enzymes XbaI and BamHI, was fused, by ligation into the two BamHI sites, to the UBILacI fragment deriving from the pMRS156 plasmid. The resulting fragment was cloned into the pMRS30 expression vector. The resulting pMRS173 vector contains a DNA fragment including the UBILacI sequence, the sequence corresponding to the 631-1275 nucleotides of the EMBL sequence

J05581, and two stop codons, TGA and TAA. This fragment is reported in fig. 9.

Example 9. Plasmid pMRS172 construction.

Plasmid pMRS167 DNA was subjected to PCR reaction in the presence of the following synthetic oligonucleotides:

V9 (5 GATCGGATCCCTGGTCTGCTCTGTCTCTGCTGCGC 3)

V10 (see example 4)

The produced DNA fragment, purified and digested with the restriction enzymes XbaI and BamHI, was fused, by ligation into the two BamHI sites, to the UBILacI fragment deriving from pMRS156 plasmid. The resulting fragment was cloned into the pMRS30 expression vector. The resulting pMRS172 vector contains a DNA fragment including the UBILacI sequence, the sequence corresponding to the 1222-1497 nucleotides of the EMBL sequence J05581, and two stop codons, TGA and TAA. This fragment is reported in fig. 10.

Example 10. Plasmid pMRS176 construction.

Plasmid pMRS167 DNA was subjected PCR reaction in the presence of the following synthetic oligonucleotides:

V3 (see example 6)

V10 (see example 4)

The produced DNA fragment, purified and digested with the restriction enzymes XbaI and BamHI, was fused, by ligation into the two BamHI sites, to the UBILacI fragment deriving from pMRS156 plasmid. The resulting fragment was cloned into the pMRS30 expression vector. The resulting pMRS176 vector contains a DNA fragment including the UBILacI sequence, the sequence corresponding to the 136-1497 nucleotides of the EMBL sequence J05581, and two stop codons, TGA and TAA. This fragment is reported in fig. 11.

Example 11. Eukaryotic cell transfection and testing for transcription.

CHO (Chinese Hamster Ovary) cells were cultured in alpha MEM supplemented with ribonucleotides and deoxyribonucleotides

at transfection time.

Dendritic cells were obtained from CD34+ hemopoietic precursors cultured in IMDM without serum, supplemented with GM-CSF, IL4, SCF, Flt3 and TNFalpha. After 7 days the obtained cell population was transfected.

Dendritic cells were obtained from monocytes isolated from PBMC (peripheral blood mononuclear cells), cultured in RPMI supplemented with FCS, GM-CSF, and IL-4. After 7 days the obtained cell population was transfected.

In each case, about one million cells were transfected with one of the plasmids reported in examples 1 to 10. Transfection was carried out using 3 µg of plasmid DNA and 4 µl of DMRIE (Gibco) by lipofection.

After 24 hours cells were harvested, washed with PBS and lysed in order to extract the mRNA.

A mRNA aliquot was subjected to RT-PCR reaction in the presence of the oligonucleotide pair specific for the transfected DNA plasmid.

This experiment was carried out for each plasmid reported in the examples 1 to 10, using the following oligonucleotide pairs: V11/V4 for pMRS166, V12/V6 for pMRS169, V13/V8 for pMRS168, V4/V10 for pMRS167, V4/V10 for pMRS175, UBIup/V4 for pMRS171, UBIup/V6 for pMRS174, UBIup/V8 for pMRS173, UBIup/V10 for pMRS172, V14/V10 for pMRS176.

As a representative example, figure 12 reports the electrophoretic analysis of the DNA fragments obtained by RT-PCR from the mRNA of the three cell populations, transfected with the pMRS169 plasmid. In this case the oligonucleotide pair V12/V6 was used.

Example 12. In vivo study results.

In the in vivo studies, the mixtures of the four fragments and the pMRS30 plasmid (vector without insert and thus used as a negative control) were used. In order to test the occurred immunization, an ELISA test was used to show the human mucin

specific antigens.

The *in vivo* studies were conducted using human MUC1 transgenic C57BL mice. As a consequence in these animals the MUC1 protein represents a self-protein. The employed vaccination schedule consists of 3 intradermic (dorsal portion, 50 micrograms DNA for each side) administrations (at days 0, 14, 28) of 100 micrograms plasmid DNA. At day 14 after the last administration, the animals were sacrificed and sera were tested for anti-human mucin antibodies.

The assayed fragment mixes, object of the present invention, stimulated a good immune response in the treated animals.

On the other hand, vaccination experiments with a 60-aminoacid peptide corresponding to the 20 aminoacids reported in fig. 2, from location 86 to location 105, repeated three times (this peptide is termed 3XTR), were also carried out.

The two vaccinations differ in the type of the elicited antibody response. The antibody titer results much more higher in the vaccination with 3XTR. Furthermore the noticed IgG subtypes are in favor of an essentially humoral (antibody) response in the case of vaccination with 3XTR, and of a cellular response (cytotoxic) in the case of vaccination with DNA. For anti-tumor therapy, a principally cytotoxic immune response is preferable. Because the experiments were carried out on transgenic mice, in whom the human mucin is "self", we can foresee a similar response in humans. This response could justify the use, as DNA vaccines, of the compounds of the present invention in the treatment of MUC1 overexpressing human tumors.

CLAIMS

1. Pharmaceutical composition containing one or more DNA molecules, encoding fragments of a protein overexpressed in
5 tumor cells in order to induce an antitumor Ag-specific immune response, in combination with suitable excipients and adjuvants.
2. Pharmaceutical composition according to claim 1 wherein the overexpressed protein is MUC-1.
3. Pharmaceutical composition according to claim 1 or 2
10 containing at least two DNA molecules each containing a cDNA sequence encoding a Mucin fragment (MUC-1).
4. Composition according to claim 3 containing at least three DNA molecules each containing a cDNA sequence encoding a Mucin fragment (MUC-1).
- 15 5. Composition according to claim 4 containing at least four DNA molecules each containing a cDNA sequence encoding a Mucin fragment (MUC-1).
6. Composition according to claims 3, 4 or 5 wherein the DNA sequences comprise about 200 to about 700 nucleotides, each
20 sequence being contiguous and possibly partially overlapping, from about 50 to about 150 nucleotides at the 3' and/or 5' end, to the adjacent one.
7. Pharmaceutical composition according to any claim from 2 to 6 wherein the used mixture consists of, at least, two plasmid DNA
25 molecules, each containing a DNA fragment selected from those whose sequences are described in figures 1, 2, 3, and 4.
8. Pharmaceutical composition according to claim 7 wherein the used mixture consists of the pool of plasmid DNA molecules, where each molecule contains a DNA fragment selected from those
30 whose sequences are described in figures 1, 2, 3, and 4.
9. Pharmaceutical composition according to claim 1 or 2 wherein a plasmid DNA molecule containing the sequence described in figure 5 is used.
10. Pharmaceutical composition according to claims 7, 8, or 9

wherein the used plasmid DNA molecules derive from the fusion of the pMRS30 expression vector in Fig. 13 to each sequence described in figures 1, 2, 3, 4, 5.

11. Pharmaceutical composition according to claims 2 to 6 wherein the used sequences, corresponding to single fragments of the protein, are preceded in the 5' termini by the sequence described in Fig. 6 encoding the ubiquitin and a *IacI* portion from *Escherichia Coli*.
12. Pharmaceutical composition according to claim 11 wherein the mixture consists of one or more sequences deriving from joining the pMRS30 expression vector, described in Fig. 13, to a DNA sequence selected from those described in figures 7, 8, 9, and 10.
13. Pharmaceutical composition according to claim 11 wherein the mixture consists of the totality of the sequences deriving from joining the pMRS30 expression vector to a DNA sequence selected from those described in figures 7, 8, 9, and 10.
14. Pharmaceutical composition according to claim 11 wherein the mixture consists of a sequence deriving from joining the pMRS30 expression vector to the sequence described in figure 11.
15. Pharmaceutical composition according to any preceding claims, further containing a cytokine or a cytokine encoding plasmid.
16. A plasmid DNA molecule consisting of the pMRS30 expression vector joined to a DNA sequence, encoding a MUC-1 protein fragment and whose sequence is selected from the group of those described in figures 1, 2, 3, 4, and 5.
17. A DNA molecule encoding a protein MUC-1 fragment preceded in its 5' terminus by the sequence described in Fig. 6.
18. A DNA molecule according to claim 17 selected from those described in figures 7, 8, 9, 10, and 11.
19. A plasmid DNA molecule obtained by joining the pMRS expression vector to a DNA molecule selected from those of claim 17 or 18.

20. Use of DNA molecules of claims 16-19 in the preparation of a composition with anti-tumor effect.

1/19

Figure 1

1 ATGACAGGTTCTGGTCATGCAAGCTCTACCCAGGTGGAGAAAAG
1▶Met Thr Gl ySer Gl yHi sAl aSer Ser Thr ProGl yGl yGl uLys
46 GAGACTTCGGCTACCCAGAGAAGTTCAGTGCCAGCTCTACTGAG
16▶Gl uThr Ser Al aThr Gl nArgSer Ser Val ProSer Ser Thr Gl u
91 AAGAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGCCACAGC
31▶LysAsnAl aVal Ser Met Thr Ser Ser Val LeuSer Ser Hi sSer
136 CCCGGTTCAGGCTCCTCCACCACTCAGGGACAGGATGCTCACTCTG
46▶ProGl ySer Gl ySer Ser Thr Thr Gl nGl yGl nAspVal Thr Leu
181 GCCCCGCCACGGAACAGCTTCAGGTTGATAA
61▶Al aProAl aThr Gl uProAl aSer Gl y.....

2/19

Figure 2

1 ATGGTGGCCAGCTCTACTGAGAAGAATGCTGTGAGTATGACCAGC
 1 Met Val Pro Ser Ser Thr Glu Lys Asn Ala Val Ser Met Thr Ser
 46 AGCGTACTCTCCAGCCACAGCCCGGTTCCAGGCTCCTCCACCACT
 16 Ser Val Leu Ser Ser His Ser Pro Glu Ser Glu Ser Ser Thr Thr
 91 CAGGGACAGGATGTCACTCTGGCCCCCGCCACGGAACCAAGCTTCA
 31 Glu Gly Glu Asp Val Thr Leu Ala Pro Ala Thr Glu Pro Ala Ser
 136 GGTTCAGCTGCCACCTGGGGACAGGATGTCACTCGGTCCCAAGTC
 46 Glu Ser Ala Ala Thr Trp Glu Glu Asp Val Thr Ser Val Pro Val
 181 ACCAGGCCAGCCCTGGGCTCCACCACCCCGCCAGCCCAAGATGTC
 61 Thr Arg Pro Ala Leu Glu Ser Thr Thr Pro Pro Ala His Asp Val
 226 ACCTCAGCCCGGACACACAGCCAGCCCGGGAAGTACTGTCCA
 76 Thr Ser Ala Pro Asp Asn Lys Pro Ala Pro Glu Ser Thr Ala Pro
 271 CCAGCACACGGTGTACCTCGGCTCCGGATACCAAGCCGGCCCCA
 91 Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg Pro Ala Pro
 316 GGTAGTACCGCCCTCCTGCCCATGGTGTCACATCTGCCCGGAC
 106 Glu Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp
 361 AACAGGCTGCAATGGGTAGTACAGCACCGCCAGTACACACGTT
 121 Asn Arg Pro Ala Leu Glu Ser Thr Ala Pro Pro Val His Asn Val
 406 ACTAGTGCTCAGGCTCTGCTAGCGGCTCAGCTTCTACTCTGGTG
 136 Thr Ser Ala Ser Glu Ser Ala Ser Glu Ser Ala Ser Thr Leu Val
 451 CACAACGGCACCTCTGCGCGCGGACACCAACCCAGCAGGCAAG
 151 His Asn Glu Thr Ser Ala Arg Ala Thr Thr Thr Pro Ala Ser Lys
 496 AGCACTCCATTCTCAATTCCCAGCTGATAA
 166 Ser Thr Pro Phe Ser Ile Pro Ser

3/19

Figure 3

1 ATGGGCTCAGCTTCTACTCTGGTGCACAACGGCACCTCTGCCAGG
 1▶ Met Gly Ser Ala Ser Thr Leu Val His Asn Gly Thr Ser Ala Arg
 46 GCTACCACAACCCAGCCAGCAAGAGCACTCCATTCTCAATTCCC
 16▶ Ala Thr Thr Thr Pro Ala Ser Lys Ser Thr Pro Phe Ser Ile Pro
 91 AGCCACCACTCTGATACTCCTACCACCTTGCCAGCCATAGCACC
 31▶ Ser His His Ser Asp Thr Pro Thr Thr Leu Ala Ser His Ser Thr
 136 AAGACTGATGCCAGTAGCACTACCATAGCACGGTACCTCCTCTC
 46▶ Lys Thr Asp Ala Ser Ser Thr His His Ser Thr Val Pro Pro Leu
 181 ACCTCCTCCAATCACAGCACTTCTCCCCAGTTGTCTACTGGGGTC
 61▶ Thr Ser Ser Asn His Ser Thr Ser Pro Gl n Leu Ser Thr Gly Val
 226 TCTTTCTTTTCTCTGTCTTTTCACATTTCAAACCTCCAGTTTAAT
 76▶ Ser Phe Phe Phe Leu Ser Phe His Ile Ser Asn Leu Gl n Phe Asn
 271 TCCTCTCTGGAGATCCCAGCACCGACTACTACCAAGAGCTGCAG
 91▶ Ser Ser Leu Gl u Asp Pro Ser Thr Asp Tyr Tyr Gl n Gl u Leu Gl n
 316 AGAGACATTTCTGAAATGTTTTTGACAGATTTATAAACAGGGGGT
 106▶ Arg Asp Ile Ser Gl u Met Phe Leu Gl n Ile Tyr Lys Gl n Gl y Gl y
 361 TTTCTGGGCCTCTCCAATATTAAGTTCAGGCCAGGATCTGTGGTG
 121▶ Phe Leu Gl y Leu Ser Asn Ile Lys Phe Arg Pro Gl y Ser Val Val
 406 GTACAATTGACTCTGGCCTTCCGAGAAGGTACCATCAATGTCAC
 136▶ Val Gl n Leu Thr Leu Ala Phe Arg Gl u Gl y Thr Ile Asn Val His
 451 GACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCAGCCTCT
 151▶ Asp Val Gl u Thr Gl n Phe Asn Gl n Tyr Lys Thr Gl u Ala Ala Ser
 496 CGATATAACCTGACGATCTCAGACGTCACGGTGAGTGATGTGCCA
 166▶ Arg Tyr Asn Leu Thr Ile Ser Asp Val Ser Val Ser Asp Val Pro
 541 TTTCTTTCTCTGCCAGTCTGGGGCTGGGGTGCCAGGCTGGGGC
 181▶ Phe Pro Phe Ser Ala Gl n Ser Gl y Ala Gl y Val Pro Gl y Trp Gl y
 585 ATCGCGCTGCTGGTCTGTCTGTCTGGTTGCGCTGGCCATT
 196▶ Ile Ala Leu Leu Val Leu Val Cys Val Leu Val Ala Leu Ala Ile
 631 GTCTATCTCATTGCTTGTGTATAA
 211▶ Val Tyr Leu Ile Ala Leu.....

4/19

Figure 4

1 ATGCTGGTGTGGTCTGTGTCTCTGGTTGCGCTGGCCATTGTCTAT
1▶MetLeuValLeuValCysValLeuValAlaLeuAlaIleValTyr
46 CTCATTGCCTTGGCTGTCTGTCTCAGTGCCGCCGAAAGAACACGGG
16▶LeuIleAlaLeuAlaValCysGlnCysArgArgLysAsnTyrGly
91 CAGCTGGACATCTTTCCAGCCCGGATACCTACCATCCTATGAGC
31▶GlnLeuAspIlePheProAlaArgAspThrTyrHisProMetSer
136 GAGTACCCACCTACCACACCCATGGGCGCTATGTGCCCCCTAGC
46▶GluTyrProThrTyrHisThrHisGlyArgTyrValProProSer
181 AGTACCGATCGTAGCCCCCTATGAGAAGGTTTCTGCAGGTAATGGT
61▶SerThrAspArgSerProTyrGluLysValSerAlaGlyAsnGly
226 GGCAGCAGCCTCTCTTACACAAACCCAGCAGTGGCAGCCACTTCT
76▶GlySerSerLeuSerTyrThrAsnProAlaValAlaAlaThrSer
271 GCCAACTTGTGATAA
91▶AlaAsnLeu•••••

Figure 5

1 ATGACAGGTTCTGGTCATGCAAGCTCTACCCCAGGTGGAGAAAAG
 1▶ Met Thr Gly Ser Gly His Ala Ser Ser Thr Pro Gly Gly Gly Lys
 46 GAGACTTCGGCTACCCAGAGAAGTTCAGTGCCAGCTCTACTGAG
 16▶ Glu Thr Ser Ala Thr Glu Arg Ser Ser Val Pro Ser Ser Thr Glu
 91 AAGAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGCCACAGC
 31▶ Lys Asn Ala Val Ser Met Thr Ser Ser Val Leu Ser Ser His Ser
 136 CCCGGTTCAGGCTCCTCCACCCTCAGGGACAGGATGTCCTCTG
 46▶ Pro Gly Ser Gly Ser Ser Thr Thr Glu Glu Gly Asp Val Thr Leu
 181 GCCCCGGCCACGGGAACCAAGCTTCAGGTTTCAGCTGCCACCTGGGGGA
 61▶ Ala Pro Ala Thr Glu Pro Ala Ser Gly Ser Ala Ala Thr Trp Gly
 226 CAGGATGTCACCTCGGTCCCACTCACCAGGCGCAGCCCTGGGCTCC
 76▶ Glu Asp Val Thr Ser Val Pro Val Thr Arg Pro Ala Leu Gly Ser
 271 ACCACCCCGCCAGCCACGATGTCACTTCAGCCCGGACAAACAG
 91▶ Thr Thr Pro Pro Ala His Asp Val Thr Ser Ala Pro Asp Asn Lys
 316 CCAGCCCGGGGAAGTACCGCTCCACCAGCACAGGTGTTCCTCG
 106▶ Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser
 361 GCTCCGGATACCAAGGCCCGGCCAGGTAGTACCGCCCTCTCTGCC
 121▶ Ala Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala
 406 CATGGTGTCACTCTGCCCCGGACACAGGCGCTGCATTGGGTAGT
 136▶ His Gly Val Thr Ser Ala Pro Asp Asn Arg Pro Ala Leu Gly Ser
 451 ACAGCACCGCCAGTACACAACGTTACTAGTGCCTCAGGCTCTGCT
 151▶ Thr Ala Pro Pro Val His Asn Val Thr Ser Ala Ser Gly Ser Ala
 496 AGCGGCTCAGCTTCTACTCTGGTGCACACGGCACCTCTGCGCGC
 166▶ Ser Gly Ser Ala Ser Thr Leu Val His Asn Gly Thr Ser Ala Arg
 541 GCGACCACAAACCCAGCGAGCAAGAGCACTCCATTCTCAATTGCC
 181▶ Ala Thr Thr Thr Pro Ala Ser Lys Ser Thr Pro Phe Ser Ile Pro
 586 AGCCCACTCTGATACTCTACCACCCCTGGCCAGCCATAGCACC
 196▶ Ser His His Ser Asp Thr Pro Thr Thr Leu Ala Ser His Ser Thr
 631 AAGACTGATGCCAGTAGCACTCACCATAGCACGGTACCTCTCTCTC
 211▶ Lys Thr Asp Ala Ser Ser Thr His His Ser Thr Val Pro Pro Leu
 676 ACCTCTCCAATCACAGCACTTCTCCCACTGTGTCTACTGGGGTC
 226▶ Thr Ser Ser Asn His Ser Thr Ser Pro Glu Leu Ser Thr Gly Val
 721 TCTTCTCTTTTCTCTCTCTTTTTCACATTTCAAACCTCCAGTTTAAAT
 241▶ Ser Phe Phe Phe Leu Ser Phe His Ile Ser Asn Leu Glu Phe Asn
 766 TCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAGCTGCAG
 256▶ Ser Ser Leu Glu Asp Pro Ser Thr Asp Tyr Tyr Glu Glu Leu Glu
 811 AGAGACATTTCTGAAATGTTTTTGACAGATTATATAACAGGGGGT
 271▶ Arg Asp Ile Ser Glu Met Phe Leu Glu Ile Tyr Lys Glu Gly
 856 TTTCTGGGCTCTCCAATATTAGTTCAGGCCAGGATCTGTGGTG
 286▶ Phe Leu Gly Leu Ser Asn Ile Lys Phe Arg Pro Gly Ser Val Val

(Continued)

6/19

Figure 5 (continued)

901 GTACAATTGACTCTGGCCTTCCGAGAAGGTACCATCAATGTCAC
301▶ Val Gl nLeuThr LeuAl aPheArg Gl uGlyThr l l eAsnVal l His
946 GACGTGGAGACACAGTTCAATCAGTATAAAACCGAAGCAGCCTCT
316▶ AspVal Gl uThr Gl nPheAsn Gl nTyrLysThr Gl uAl aAl aSer
991 CGATATAACCTGACGATCTCAGACGTCAGCGTGAGTGATGTGCCA
331▶ ArgTyrAsnLeuThr l l eSerAspVal Ser Val SerAspVal Pro
1036 TTTCTTTTCTCTGCCAGTCTGGGGCTGGGGTGCCAGGCTGGGGC
346▶ PheProPheSer Al aGl nSer Gl yAl aGlyVal ProGlyTyrGly
1081 ATCGCGCTGCTGGTGCTGGTCTGTGTTCTGGTTCGGCTGGCCATT
361▶ l l eAl aLeuLeuVal LeuVal CysVal LeuValAl aLeuAl l e
1126 GTCTATCTCATTTGCCCTTGGCTGTCTGTCTGTCAGTGCCGCCGAAGAAC
376▶ ValTyrLeu l l eAl aLeuAl aVal CysGl nCysArgArgLysAsn
1171 TACGGGCAGCTGGACATCTTCCAGCCCGGGATACCTACCATCCT
391▶ TyrGly Gl nLeuAsp l l ePheProAl aArgAspThr TyrHisPro
1216 ATGAGCGAGTACCCACCTACCACACCCATGGGCGCTATGTGCC
406▶ MetSer Gl uTyrProThr TyrHisThrHisGlyArgTyrVal Pro
1261 CCTAGCAGTACCGATCGTAGCCCTATGAGAAGGTTTCTGCAGGT
421▶ ProSerSer ThrAspArgSer ProTyrGl uLysVal SerAl aGly
1306 AATGGTGGCAGCAGCCTCTCTACACAAACCCAGCAGTGGCAGCC
436▶ AsnGl yGlySer Ser LeuSer TyrThrAsnProAl aValAl aAl a
1351 ACTTCTGCCAATGTGTGATAA
451▶ ThrSerAl aAsnLeu*****

7/19

Figure 6

1 ATGCAGATCTTGGTGAAGACCCCTGACTGGTAAGACCATCACTCTC
1▶ Met Gl n l l e Phe Val Lys Thr Leu Thr Gl y Lys Thr l l e Thr Leu
46 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAGATC
16▶ Gl u Val Gl u Pro Ser Asp Thr l l e Gl u Asn Val Lys Al a Lys l l e
91 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
31▶ Gl n Asp Lys Gl u Gl y l l e Pro Pro Asp Gl n Gl n Arg Leu l l e Phe
136 GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
46▶ Al a Gl y Lys Gl n Leu Gl u Asp Gl y A rg Thr Leu Ser Asp Tyr Asn
181 ATCCAGAAAGAGTCCACCTGCACCTGGTGTCTCCGTCTCAGAGGT
61▶ l l e Gl n Lys Gl u Ser Thr Leu Hi s Leu Val Leu Arg Leu Arg Gl y
226 GGGAGGCACGGTAGTGGTGCATGGCTGTGGCCGCTCTCGCTGGTG
76▶ Gl y A rg Hi s Gl y Ser Gl y Al a Trp Leu Leu Pro Val Ser Leu Val
271 AAAAGAAAAACACCCCTGGCGCCCAATACGCAAACCGCCTCTCCC
91▶ Lys Arg Lys Thr Thr Leu Al a Pro Asn Thr Gl n Thr Al a Ser Pro
316 CGCGCGTTGGCCGATTCTTAATGCAGCTGGCAGCAGAGTTTCC
106▶ A rg Al a Leu Al a Asp Ser Leu Met Gl n Leu Al a Arg Gl n Val Ser
361 CGAGGATCC
121▶ A rg Gl y Ser

8/19

Figure 7

1 ATGCAGATCTTCGTGAAGACCTGACTGGTAAGACCATCACTCTC
1 Met Gl n I l e Phe Val Lys Thr Leu Thr Gly Lys Thr I l e Thr Leu
46 GAAGTGGAGCCGAGTGACACCATGAGAATGTCAAGGCAAAGATC
16 Gl u Val Gl u Pro Ser Asp Thr I l e Gl u Asn Val Lys Ala Lys I l e
91 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
31 Gl n Asp Lys Gl u Gl y I l e Pro Pro Asp Gl n Gl n Arg Leu I l e Phe
136 GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
46 Al a Gl y Lys Gl n Leu Gl u Asp Gl y A rg Thr Leu Ser Asp Tyr Asn
181 ATCCAGAAAGAGTCCACCCCTGCACCTGGTGTCTCCGTCTCAGAGGT
61 I l e Gl n Lys Gl u Ser Thr Leu Hi s Leu Val Leu Arg Leu Arg Gly
226 GGGAGGCACGGTAGTGGTGCATGGCTGTGCCCCGTCTCGCTGGTG
76 Gl y A rg Hi s Gl y Ser Gl y Ala Trp Leu Leu Pro Val Ser Leu Val
271 AAAAGAAAAACCCCTGGCGCCCAATACGCAACCGCCTCTCCCC
91 Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Gl n Thr Ala Ser Pro
316 CGCGCGTTGGCCGATTTCATTAATGCAGCTGGCAGCAGAGTTTCC
106 A rg Ala Leu Ala Asp Ser Leu Met Gl n Leu Ala Arg Gl n Val Ser
361 CGAGGATCCACAGGTTCTGGTCATGCAAGCTCTACCCAGGTGGA
121 A rg Gl y Ser Thr Gl y Ser Gl y Hi s Ala Ser Ser Thr Pro Gl y Gl y
406 GAAAAGGAGACTTCGGCTACCCAGAGAAGTTCAAGTGGCCAGCTCT
136 Gl u Lys Gl u Thr Ser Ala Thr Gl n Arg Ser Ser Val Pro Ser Ser
451 ACTGAGAAGAAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGC
151 Thr Gl u Lys Asn Ala Val Ser Met Thr Ser Ser Val Leu Ser Ser
496 CACAGCCCCGTTTCAGGCTCCTCCACCCTCAGGACAGGATGTC
166 Hi s Ser Pro Gl y Ser Gl y Ser Ser Thr Thr Gl n Gl y Gl n Asp Val
541 ACTCTGGCCCCGGCCACGGAACCAAGCTTCAGGTTGATAA
181 Thr Leu Ala Pro Ala Thr Gl u Pro Ala Ser Gl y •••••

9/19

Figure 8

1 ATGCAGATCTTCTGTGAAGACCCTGACTGGTAAGACCATCACTCTC
 1 Met Gl n I l e Phe Val Lys Thr Leu Thr Gly Lys Thr I l e Thr Leu
 46 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAGATC
 16 Gl u Val Gl u Pro Ser Asp Thr I l e Gl u Asn Val Lys Ala Lys I l e
 91 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31 Gl n Asp Lys Gl u Gly I l e Pro Pro Asp Gl n Gl n Arg Leu I l e Phe
 136 GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46 A l a Gl y Lys Gl n Leu Gl u Asp Gl y A rg Thr Leu Ser Asp Tyr Asn
 181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGTCTCGTCTCAGAGGT
 61 I l e Gl n Lys Gl u Ser Thr Leu Hi s Leu Val Leu Arg Leu Arg Gl y
 226 GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76 Gl y A rg Hi s Gl y Ser Gl y Ala T rp Leu Leu Pro Val Ser Leu Val
 271 AAAAGAAAACCACCCTGCCGCCCAATACGCAAAACCGCCTCTCCC
 91 Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Gl n Thr Ala Ser Pro
 316 CGCGCGTTGGCCGATTCATTAAATGCAGCTGGCAGCAGAGTTTCC
 106 A rg Ala Leu Ala Asp Ser Leu Met Gl n Leu Ala Arg Gl n Val Ser
 361 CGAGGATCCGTGCCAGCTCTACTGAGAAGAATGCTGTGAGTATG
 121 A rg Gl y Ser Val Pro Ser Ser Thr Gl u Lys Asn Ala Val Ser Met
 406 ACCAGCAGCGTACTCTCCAGCCACAGCCCCGGTTCAGGCTCTCTCC
 136 Thr Ser Ser Val Leu Ser Ser Hi s Ser Pro Gl y Ser Gl y Ser Ser
 451 ACCACTCAGGGACAGGATGTCACTCTGGCCCCGGCCACGGAAACCA
 151 Thr Thr Gl n Gl y Gl n Asp Val Thr Leu Ala Pro Ala Thr Gl u Pro
 496 GCTTCAGGTTTCAGCTGCCACCTGGGGACAGGATGTCACCTCGGTC
 166 A Ser Gl y Ser Ala Ala Thr T rp Gl y Gl n Asp Val Thr Ser Val
 541 CCAGTCACCAAGGCCAGCCCTGGGCTCCACCACCCCGCCAGCCAC
 181 Pro Val Thr A rg Pro Ala Leu Gl y Ser Thr Thr Pro Pro Ala Hi s
 586 GATGTCACTCAGCCCCGGACAACAAGCCAGCCCCGGGAAGTACT
 196 Asp Val Thr Ser Ala Pro Asp Asn Lys Pro Ala Pro Gl y Ser Thr
 631 GCTCCACCAGCACACGGTGTACCTCGGCTCCGGATACCAGGCCG
 211 Ala Pro Pro Ala Hi s Gl y Val Thr Ser Ala Pro Asp Thr A rg Pro
 676 GCCCCAGGTAGTACCGCCCCCTCTGCCCATGGTGTACATCTGCC
 226 Ala Pro Gl y Ser Thr Ala Pro Pro Ala Hi s Gl y Val Thr Ser Ala
 721 CCGGACAACAGCCCTGCATTGGGTAGTACAGCACCGCCAGTACAC
 241 Pro Asp Asn Arg Pro Ala Leu Gl y Ser Thr Ala Pro Pro Val Hi s
 766 AACGTTACTAGTGCTCAGGCTCTGCTAGCGGCTCAGCTTCTACT
 256 Thr Asn Val Thr Ser Ala Ser Gl y Ser Ala Ser Gl y Ser Ala Ser Thr
 811 CTGGTGCACAACGGCACCTCTGCGCGCGGACCACAACCCAGCG
 271 Leu Val Hi s Asn Gl y Thr Ser Ala Arg Ala Thr Thr Thr Pro Ala
 856 AGCAAGAGCACTCCATTCTCAATTCCCAGCTGATAA
 286 Ser Lys Ser Thr Pro Phe Ser I l e Pro Ser

Figure 9

1 ATGCAGATCTTCGTGAAGACCCCTGACTGGTAAGACCATCACTCTC
 1 Met Gl n l l e P h e V a l L y s T h r L e u T h r G l y L y s T h r l l e T h r L e u
 46 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC
 16 Gl u V a l G l u P r o S e r A s p T h r l l e G l u A s n V a l L y s A l a L y s l l e
 91 CAAGACAAGGAAGGCATCCCTCTGACCAGCAGAGGCTCATCTTT
 31 Gl n A s p L y s G l u G l y l l e P r o P r o A s p G l n G l n A r g L e u l l e P h e
 136 GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTCTGACTACAAC
 46 A l a G l y L y s G l n L e u G l u A s p G l y A r g T h r L e u S e r A s p T y r A s n
 181 ATCCAGAAAGAGTCCACCCCTGCACCTGGTGTCTCGTCTCAGAGGT
 61 l l e G l n L y s G l u S e r T h r L e u H i s L e u V a l L e u A r g L e u A r g G l y
 226 GGGAGGCACGGTAGTGGTGCATGGCTGTGGCCGCTCTCGCTGGTG
 76 G l y A r g H i s G l y S e r G l y A l a T r p L e u L e u P r o V a l S e r L e u V a l
 271 AAAAGAAAAACCCCTTGGCGCCCAATACGCAAAACCGCTCTCC
 91 L y s A r g L y s T h r T h r L e u A l a P r o A s n T h r G l n T h r A l a S e r P r o
 316 CGCGCGTGGCCGATTCATTATGACAGCTGGCAGCAGAGGTTTCC
 106 A r g A l a L e u A l a A s p S e r L e u M e t G l n L e u A l a A r g G l n V a l S e r
 361 CGAGGATCCGGCTCAGCTTCTACTCTGGTGCACAACGGCACCTCT
 121 A r g G l y S e r G l y S e r A l a S e r T h r L e u V a l H i s A s n G l y T h r S e r
 406 GCCAGGCTTACCACAACCCAGCCAGCAGCAGCACTCCATTCTCA
 136 A l a A r g A l a T h r T h r T h r P r o A l a S e r L y s S e r T h r P r o P h e S e r
 451 ATTCCAGCCACCACTCTGATACTCTACCACCCCTTCCAGCCAT
 151 l l e P r o S e r H i s H i s S e r A s p T h r P r o T h r T h r L e u A l a S e r H i s
 496 AGCACCAGACTGATGCCAGTAGCACTCACCATAGCAGCGTACCT
 166 S e r T h r L y s T h r A s p A l a S e r S e r T h r H i s H i s S e r T h r V a l P r o
 541 CCTCTCACCTCCTCCCAATCACAGCACTTCTCCCCAGTTGTCTACT
 181 P r o L e u T h r S e r S e r A s n H i s S e r T h r S e r P r o G l n L e u S e r T h r
 586 GGGTCTCTTCTTTTCTCTGCTCTTTTCACATTCAAACTCCAG
 196 G l y V a l S e r P h e P h e P h e L e u S e r P h e H i s l l e S e r A s n L e u G l n
 631 TTTAATTCCTCTCTGGAAGATCCAGCAGCCGACTACTACCAAGAG
 211 P h e A s n S e r S e r L e u G l u A s p P r o S e r T h r A s p T y r T y r G l n G l u
 676 CTGCAGAGAGACATTCTGAAATGTTTTGAGATTATATAACAA
 226 L e u G l n A r g A s p l l e S e r G l u M e t P h e L e u G l n l l e T y r L y s G l n
 721 GGGGTTTCTTGGGCTCTCCCAATATTAAAGTTCAGGCCAGGATCT
 241 G l y G l y P h e L e u G l y L e u S e r A s n l l e L y s P h e A r g P r o G l y S e r
 756 GTGGTGGTACAATTGACTCTGGCCCTCCGAGAAGGTACCATCAAT
 256 V a l V a l V a l G l n L e u T h r L e u A l a P h e A r g G l u G l y T h r l l e A s n
 811 GTCCACGACGTGAGACACAGTTCAATCAGTATAAAACCGGAAGCA
 271 V a l H i s A s p V a l G l u T h r G l n P h e A s n G l n T y r L y s T h r G l u A l a
 856 GCTCTCGATATAACCTGACGATCTCAGACGTCAGCGTGAGTGAT
 286 A l a S e r A r g T y r A s n L e u T h r l l e S e r A s p V a l S e r V a l S e r A s p
 901 GTGCCATTCTCTTCTTCTGCCAGTCTGGGGCTGGGTGCCAGGC
 301 V a l P r o P h e P r o P h e S e r A l a G l n S e r G l y A l a G l y V a l P r o G l y
 946 TGGGGCATCCGCTGCTGGTGTCTGGTCTGTCTTCTGGTTCGCGCTG
 316 T r p G l y l l e A l a L e u L e u V a l L e u V a l C y s V a l L e u V a l A l a L e u
 991 GGCATTGTCTATCTCATTGCCCTGTGATAA
 331 A l a l l e V a l T y r L e u l l e A l a L e u

11/19

Figure 10

1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC
 1▶ Met Gl n I l e Phe Val Lys Thr Leu Thr Gl y Lys Thr I l e Thr Leu
 46 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAGATC
 16▶ Gl u Val Gl u Pro Ser Asp Thr I l e Gl u Asn Val Lys Al a Lys I l e
 91 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ Gl n Asp Lys Gl u Gl y I l e Pro Pro Asp Gl n Gl n Arg Leu I l e Phe
 136 GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ Al a Gl y Lys Gl n Leu Gl u Asp Gl y A rg Thr Leu Ser Asp Tyr Asn
 181 ATCCAGAAAGAGTCCACCCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ I l e Gl n Lys Gl u Ser Thr Leu Hi s Leu Val Leu Arg Leu Arg Gl y
 226 GCGAGGCACGCTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ Gl y A rg Hi s Gl y Ser Gl y Al a Trp Leu Leu Pro Val Ser Leu Val
 271 AAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
 91▶ Lys Arg Lys Thr Thr Leu Al a Pro Asn Thr Gl n Thr Al a Ser Pro
 316 CGCGCGTTGGCCGATTCAITTAATGCAGCTGGCAGCAGGTTTCC
 106▶ A rg Al a Leu Al a Asp Ser Leu Met Gl n Leu Al a Arg Gl n Val Ser
 361 CGAGGATCCCTGGTCTGGTCTGTGTCTGGTTGCGCTGGCCATT
 121▶ A rg Gl y S er Leu Val Leu Val Cys Val Leu Val Al a Leu Al a I l e
 406 GTCTATCTCATTTGCCCTTGGCTGTCTGTCTAGTGCCGCCGAAAGAAC
 136▶ Val Tyr Leu I l e Al a Leu Al a Val Cys Gl n Cys Arg Arg Lys Asn
 451 TACGGGCAGCTGGACATCTTTCCAGCCCGGATACCTACCATCCT
 151▶ Tyr Gl y Gl n Leu Asp I l e Phe Pro Al a Arg Asp Thr Tyr Hi s Pro
 496 ATGAGCGAGTACCCACCTTACCACCCCATGGCGCTATGTGCC
 166▶ Met Ser Gl u Tyr Pro Thr Tyr Hi s Thr Hi s Gl y A rg Tyr Val Pro
 541 CCTAGCAGTACCGATCGTAGCCCTATGAGAAGGTTTCTGCAGGT
 181▶ Pro Ser Ser Thr Asp Arg Ser Pro Tyr Gl u Lys Val Ser Al a Gl y
 586 AATGGTGGCAGCAGCCTCTCTTACACAAACCCAGCAGTGGCAGCC
 196▶ Asn Gl y Gl y Ser Ser Leu Ser Tyr Thr Asn Pro Al a Val Al a Al a
 631 ACTTCTGCCAACTTGTGATAA
 211▶ Thr Ser Al a Asn Leu.....

12/19

Figure 11

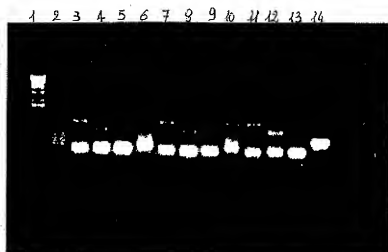
1 ATGCAGATCTTCGTGAAGACCTGACTGGTAAGACCATCACTCTC
 1▶ Met Gl n I l e P h e V a l L y s T h r L e u T h r G l y L y s T h r I l e T h r L e u
 46 GAAGTGGAGCCGAGTGACACCATTTGAGAATGTCAAGGCAAAGATC
 16▶ Gl u V a l G l u P r o S e r A s p T h r I l e G l u A s n V a l L y s A l a L y s I l e
 91 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ Gl n A s p L y s G l u G l y I l e P r o P r o A s p G l n G l n A r g L e u I l e P h e
 136 GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ A l a G l y L y s G l n L e u G l u A s p G l y A r g T h r L e u S e r A s p T y r A s n
 181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ I l e G l n L y s G l u S e r T h r L e u H i s L e u V a l L e u A r g L e u A r g G l y
 226 GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ G l y A r g H i s G l y S e r G l y A l a T r p L e u L e u P r o V a l S e r L e u V a l
 271 AAAAGAAAAACCACCCTGGCGCCCAATACGCAAAACCGCTCTCCC
 91▶ L y s A r g L y s T h r T h r L e u A l a P r o A s n T h r G l n T h r A l a S e r P r o
 316 CGCGCGTTGGCCGATTCTTAATGCAGCTGGCAGCAGAGGTTTCC
 106▶ A r g A l a L e u A l a A s p S e r L e u M e t G l n L e u A l a A r g G l n V a l S e r
 361 CGAGGATCCACAGGTTCTGGTCATGCAAGCTCTACCCAGGTGGA
 121▶ A r g G l y S e r T h r G l y S e r G l y H i s A l a S e r S e r T h r P r o G l y G l y
 406 GAAAAGGAGACTTCGGCTACCCAGAGAAGTTTCAGTGCCCGCTCT
 136▶ G l u L y s G l u T h r S e r A l a T h r G l n A r g S e r S e r V a l P r o S e r S e r
 451 ACTGAGAAGATGCTGTGACTATGACCAGCAGCGTACTCTCAGC
 151▶ T h r G l u L y s A s n A l a V a l S e r M e t T h r S e r S e r V a l L e u S e r S e r
 496 CACAGCCCCGGTTCAGGCTCTCTCCACCCTCAGGACAGGATGTC
 166▶ H i s S e r P r o G l y S e r G l y S e r S e r T h r G l n G l y G l n A s p V a l
 541 ACTCTGGCCCCCGGCCAGGAAACAGCTTCAGGTTTCAGTGCCACC
 181▶ T h r L e u A l a P r o A l a T h r G l u P r o A l a S e r G l y S e r A l a A l a T h r
 586 TGGGGACAGGATGTCACCTCGGTCCCGTCCAGTCACCAGGCCCGCTG
 196▶ T r p G l y G l n A s p V a l T h r S e r V a l P r o V a l T h r A r g P r o A l a L e u
 631 GGCTCCACCACCCCGCCAGCCACGATGTCACTCAGCCCCCGGAC
 211▶ G l y S e r T h r T h r P r o P r o A l a H i s A s p V a l T h r S e r A l a P r o A s p
 676 ACAAGCCAGCCCCCGGAAGTACCGCTCCACCAGCACACGGTGTT
 226▶ A s n L y s P r o A l a P r o G l y S e r T h r A l a P r o P r o A l a H i s G l y V a l
 721 ACCTCGGCTCCGGATACAGGCCCGGCCAGGTAGTACCGCCCT
 241▶ T h r S e r A l a P r o A s p T h r A r g P r o A l a P r o G l y S e r T h r A l a P r o
 766 CCTGCCCATGGTGTACATCTGCCCGGACAACAGGCCCTGCATTG
 256▶ P r o A l a H i s G l y V a l T h r S e r A l a P r o A s p A s n A r g P r o A l a L e u
 811 GGTAGTACAGCACCGCCAGTACACAACGTTACTAGTGCCTCAGGC
 271▶ G l y S e r T h r A l a P r o P r o V a l H i s A s n V a l T h r S e r A l a S e r G l y
 856 TCTGCTAGCGGCTCAGCTTCTACTCTGGTGCAACAGGCACCTCT
 286▶ S e r A l a S e r G l y S e r A l a S e r T h r L e u V a l H i s A s n G l y T h r S e r

(Continued)

13/19

Figure 11 (continued)

901 GCGCGCGGACCACAAACCCAGCGAGCAAGAGCACTCCATTCTCA
301▶AlaArgAlaThrThrThrProAlaSerLysSerThrProPheSer
946 ATPTCCAGCCCACTCTGATACTCCTACCACCTTGCCAGGCAT
316▶IleProSerHisHisSerAspThrProThrThrLeuAlaSerHis
991 AGCACCAAGACTGATGCCAGTAGCACTCACCATAGCACGGTACCT
331▶SerThrLysThrAspAlaSerSerThrHisHisSerThrValPro
1036 CCTCTCACCTCCTCAATCAGCACTTCTCCCACTTGTCTACT
346▶ProLeuThrSerSerAsnHisSerThrSerProGlnLeuSerThr
1081 GGGGTCTCTTTCTTTTCTCTCTTTTCACATTTCAAACCTCCAG
361▶GlyValSerPhePhePheLeuSerPheHsIleSerAsnLeuGln
1126 TTTAATTCTCTCTGGAAGATCCCAAGCAGCACTACTACCAAGAG
376▶PheAsnSerSerLeuGluAspProSerThrAspTyrTyrGlnGlu
1171 CTGCAGAGAGACATTCTGAAATGTTTTGCAGATTTATAAACAA
391▶LeuGlnArgAspIleSerGluMetPheLeuGlnIleTyrLysGln
1216 GGGGGTTTCTGGGCCTCTCCAATTTAAGTTCAGGCCAGGATCT
406▶GlyGlyPheLeuGlyLeuSerAsnIleLysPheArgProGlySer
1261 GTGGTGGTACAATTGACTCTGGCCTTCGAGAAGGTACCATCAAT
421▶ValValValGlnLeuThrLeuAlaPheArgGluGlyThrIleAsn
1306 GTCCACGACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCA
436▶ValHisAspValGluThrGlnPheAsnGlnTyrLysThrGluAla
1351 GCCTCTCGATATAACCTGACGATCTCAGACGTCAGCGTGAGTGAT
451▶AlaSerArgTyrAsnLeuThrIleSerAspValSerValSerAsp
1396 GTGCCATTCTCTTCTCTGCCCAGCTCTGGGGCTGGGGTGCCAGGC
466▶ValProPheProPheSerAlaGlnSerGlyAlaGlyValProGly
1441 TGGGGCATCGCGCTGCTGGTCTGTGTTCTGGTTGGCGCTG
481▶TrpGlyIleAlaLeuLeuValLeuValCysValLeuValAlaLeu
1486 GCCATTGTCTATCTCATTGCCCTTGGCTGTCTGTCAGTCCCGCGA
496▶AlaIleValTyrLeuIleAlaLeuAlaValCysGlnCysArgArg
1531 AAGAACTACGGGAGCTGGACATCTTCCAGCCCGGGATACCTAC
511▶LysAsnTyrGlyGlnLeuAspIlePheProAlaArgAspThrTyr
1576 CATCTATGAGCGAGTACCCCACTACCACACCATGGGCGCTAT
526▶HisProMetSerGluTyrProThrTyrHisThrHisGlyArgTyr
1621 GTGCCCCCTAGCAGTACCGATCGTAGCCCCCTATGAGAAGGTTTCT
541▶ValProProSerSerThrAspArgSerProTyrGluLysValSer
1666 GCAGGTAATGGTGGCAGCAGCCTCTCTTACACAAACCCAGCAGTG
556▶AlaGlyAsnGlyGlySerSerLeuSerTyrThrAsnProAlaVal
1711 GCAGCCACTTCTGCCAAGTTGTGATAA
571▶AlaAlaThrSerAlaAsnLeu*****



15/19

Figure 13

1 CCAGGAAGCTCCTCTGTGTCTCTCATAAACCTTAACCTCCTCTACTTGAGA
51 GGACATTCCAATCATAGGCTGCCCATCCACCTCTGTGTCTCTCTGTATAA
101 TTAGTCACTTAACAAAAAGGAAATTGGGTAGGGGTTTTTCACAGACGCG
151 TTTCTAAGGGTAATTTTAAAAATATCTGGGAAGTCCCTTCCCATGCTGTGT
201 TCCAGAAAGTGTGTGTAACAGCCCCACAAATGTCAACAGCAGAAACATACA
251 AGCTGTCAAGTTTGACAAAGGGCCAACACCTGCTCATCAAGAAGCACT
301 GTGGTTGCTGTGTAGTAATGTGCAAAACAGGAGGCACATTTTCCCCAOC
351 TGTGTAGGTTCCAAAATATCTAGTGTTTTCATTTTACTTGGATCAGGAA
401 CCCAGCACTCCACTGGATAAGCAATTATCCTTATCCAAAACAGCCTTGTGG
451 TCAGTGTTCATCTGCTGACTGTCAACTGTAGCAITTTTTGGGGTTACAGT
501 TTGAGCAGGATATTTGGTCTGTAGTTTGTAAACACACCTTCAGCTCCA
551 AAGGTTCCCCACCAACAGCAAAAAAATGAAATTTGACCTTGAATGGGT
601 TTTCCAGCAACATTTTCATGAGTTTTTTGTGTCCCTGAATGCAAGTTTAA
651 CATAGCAGTTACCCCAATACCTCAGTTTTTAACAGTAACAGCTTCCACACA
701 TCMAATATTTCCACAGGTTAAGTCTCATTTAAATTAGGCAAGGAATT
751 CTTGAAGACGAAAGGGCCCTGTGATACGCCATTTTTATAGGTTAATGTC
801 ATGATAATAATGGTTTCTTAGACGTCAAGTGGCACTTTTCGGGGAAATGT
851 GCGCGGAACCCCATTTTGTATTATTTTCTAAATACATTCAAATATGTATC
901 CGCTCATGAGACAATAACCTTGATAAATGCTTCAATAATATTGAAAAAGG
951 AAGAGTATGAGTATTCACATTTCCGCTGTGCGCCTTATTCCTTTTTTGC
1001 GGCATTTTGCTTCTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAA

Figure 13

2151 TAGTTAGGOCACCACTTCAAGAACTCTGTAGCACOGCCTACATACCTGCG
2201 TCTGCTAATCCTGTTACCACTGGCTGCTGCCAGTGGCGATAAGTGTGTG
2251 TTACCGGGTIGGACTCAAGACGATAGTTACCGGATAAGGCGCAGGGTGG
2301 GCCTGAACGGGGGGTTCTGTGCACACAGCCAGCTTGGAGGGAACGACCTA
2351 CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAGCTTC
2401 CCGAAGGGAGAAAGCGCGCAGGTATCCGGTAAGCGCAGGGTGGGAACA
2451 GGAGAGGCGCAAGAGGGAGCTTCCAGGGGGAAAAGCGCTGTATCTTTATAG
2501 TCCTGTGGGGTTTCCGCACTCTGACTTGAGCGTGGATTTTGTGATGCT
2551 CGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGCGCTTTTAA
2601 CGGTTCTTGGCCTTTTGTGCGCCTTTTGTCTCATGTTCTTCTCGGTT
2651 ATCCCTGTATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATA
2701 CCGCTCGCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAAGGAGAA
2751 GCGGAAGAGCGCCTGATCGGGTATTTTCTCCTTACGCATCTGTGGGTAT
2801 TTCACACGCGATATGGTGCACCTCTAGTACAATCTGCTCTGATGCCGAT
2851 AGTTAAGCCAGTATACAATCAATATTGGCCATTAGCCATATTATTATTG
2901 GTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTAT
2951 CCATATCATATAATGTACATTTATATTGGCTCATGTCCAACATTACCGCC
3001 ATGTTGACATTGATTATTGACTAGTTATTAAATAGTAATCAATTACGGGT
3051 CATTAGTTATAGCCCATATATGGAGTTCCGCTTACATAACTTACGGTA
3101 AATGGCCCGCTGGCTGACCGCCCAAGACCCCGCCCATTTGACGTCAAT
3151 AATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTC
3201 AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTG
3251 TATCATATGCCAAGTAAGCCCCCTATTGACGTCATGACGGTAAATGGCC

(Continued)

Figure 13 (Continued)

3301 CGCCTGGCAATTATGCCCCAGTACATGACCTTATGGGACTTTCTACTTGGC
3351 AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG
3401 CAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAG
3451 TCTCCACCCCATGACGTCAATGGGAGTTTGGTTTGGCACCAAAATCAAC
3501 GGGACTTTCCAAAATGTGTAACTCCGCCCATTTGACGCAATGGGC
3551 GGTAGGCGGTGACGGTGGGAGTCTATATAAGCAGAGCTCGTTTAGTGA
3601 CCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTGGACCTCCATAGAA
3651 GACACGGGACCGATCCAGCCTCCGCGGCCGGGAAGGTGCATTGGAAAG
3701 CGGATTCOCCTGCCAAGAAAGCTTGTCTAGAACCGGAGAGCTCCTGA
3751 GAACCTCAGGGTGAGTTTGGGGACCTTGATGTTCTTTCTTTTCGCTA
3801 TTGTAATAATTCATGTTATATGGAGGGGCAAGTTTTCAGGGTGTGTTT
3851 AGAATGGGAAGATGTCCTTGTATCACCATGGACCTCATGATAATTTTG
3901 TTTCTTTCACCTTTCTACTCTGTGACAACCAATTGTCTCTCTTATTCTT
3951 TTTCATTTTCTGTAACTTTTTCGTAAACTTTAGCTTGCAATTTGTAACGA
4001 ATTTTAAATTCACCTTTTGTATTATTGTGAGATTGTAAGTACTTTCTCTA
4051 ATCACTTTTTTTTCAAGGCAATCAGGGTATATTATATTGTACTTCAGCAC
4101 AGTTTATAGAGAACAAATGTTTATAATTAATGATAAGGTAGAAATTTCTG
4151 CATATAAATCTGGCTGGCGTGGAAATATTCTTATTTGGTAGAAACAATA
4201 CATCTTGGTCAATCTCTGCCCTTTCTCTTTATGGTTACAATGATATACAC
4251 TGTTTGAGATGAGGATAAAATACTCTGAGTCCAAACCGGGCCCTCTGCT
4301 AACCATGTTCACTCTCTCTTTTCTTTTCTACAGCTCCTGGGCAAGGTGCT
4351 GGTGTTGTGCTGTCTCATCATTTTGGCAAGAAATCACTCCTCAGGTGC
4401 AGGCTGCCTATCAGAAGGTGGTGGCTGGTGGCCAATGCCCTGGCTCAC

(Continued)

Figure 13 (Continued)

1051 AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGAT
1101 CTCACACAGCGTAAGATCCTTGAGAGTTTTGCGCCCCGAAGACGTTTTCC
1151 AATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGCGTATATCCCGTG
1201 TTGACGCGGGCAAGAGCACTCGGTGCGCGCATACTATTCTCAGAAT
1251 GACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCAT
1301 GACAGTAAGAGAAATTATGCACTGCTGCCATAACCATGAGTGATAACACTG
1351 CGGCCAACTTACTTCTGACAAAGATCGGAGGACGGAAGAGCTAACCGCT
1401 TTTTTCACAAACATGGGGATCATGTAACCTCGCCTTGATCGTTGGGAAAC
1451 GGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCAAGATCGCTG
1501 CAGCAATGGCAACAACGTTGGCGCAACTATTAACTGGCGAACTACTTACT
1551 CTAGCTTCCCGCAACAATTAATAGACTGGATGAGGCGGATAAAGTTGC
1601 AGGACCATTCTCGGCTCGGCCCTTCGGCTGCTGTTTATGCTGATA
1651 AATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGG
1701 CCAGATGGTAAGCCCTCCCGTATGTAGTTATCTACAGACGGGGAGTCA
1751 GGCAACTATGATGAACGAATATAGACAGATCGCTGAGATAGGTGCCYCAC
1801 TGATTAAGCATTGGTAAGTCTCAGACCAAGTTTACTCATATATACTTTAG
1851 ATTGATTTAAACCTTCATTTTTAATTTTAAAGGATCAGGTGAAGATCCT
1901 TTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACT
1951 GAGCGTCAGACCCCGTAGAAAAGATCAAGGATCTTCTTGAGATCCTTTT
2001 TTTCTGCGGTAATCTGCTGCTTGCAAAACAAAAACCAACGCTACCAAGC
2051 GGTTGGTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAA
2101 CTGGCTTCAGCAGAGCGCAGATACCAATACTGCTCTTCTAGTGTAGCGG

(Continued)

19/19

Figure 13 (Continued)

4451 AAATACCACTGAGATCTTTTCCCTCTGCCAAAATTATGGGACATCAT
4501 GAAGCCCTTGAGCATCTGACTCTGGCTAATAAGGAATTATTTTCA
4551 TTGCAATAGTGTGTTGGAATTTTTGTGTCTCTCACTGGGAAGGACATAT
4601 GGGAGGGCAATCATTTAAACATCAGATGAGTATTTGGTTAGAGTTT
4651 GGCAACATATGCCATATGCTGGCTGCCATGAACAAAGGTGCTATAAAGA
4701 GGTCAATCAGTATATGAACAGCCCCCTGCTGTCCATTCTTATTCATAG
4751 AAAAGCCTTGACTTGAGGTAGATTTTTTTATATTTTGTTTTGTGTTAT
4801 TTTTTTCTTTAACATCCCTAAAATTTTCCTTACATGTTTTACTAGCCAGA
4851 TTTTCCCTCTCTCTGACTACTCCAGTCATAGCTGTCCCTCTTCTCTG
4901 GATCC

SEQUENCE LISTING

<110> MENARINI RICERCHE S.p.A.

<120> PHARMACEUTICAL COMPOSITION, CONTAINING FRAGMENTS OF AN
ANTIGENIC PROTEIN ENCODING DNA ENDOWED WITH ANTI-TUMOR
EFFECT

<130> 5653MBUR

<140>

<141>

<150> MI98A002330

<151> 1998-10-30

<160> 35

<170> PatentIn Ver. 2.1

<210> 1

<211> 213

<212> DNA

<213> human

<400> 1

atgacagggt ctggtcatgc aagctctacc ccaggtggag aaaaggagac ttcggtacc 60
cagagaagtt cagtgcccag ctctactgag aagaatgctg tgaatatgac cagcagcgt 120
ctctccagcc acagccccgg ttcagggtcc tccaccactc agggacagga tgtcactctg 180
gccccggcca cggaaccagc ttcagggtga taa 213

<210> 2

<211> 525

<212> DNA

<213> human

<400> 2

atgggtccca gctctactga gaagaatgct gtgagtatga ccagcagcgt actctccagc 60
cacagccccg gttcagggtc ctccaccact cagggacagg atgtcactct ggccccggcc 120
acggaaccag ctccagggtc agctgccacc tggggacagg atgtcacctc ggtcccagtc 180
accaggccag cctctgggtc caccaccccg ccagcccacg atgtcacctc agccccggag 240
aacaagccag cccccgggaag tactgtctcc cagcacacag gtgttacctc ggctccggat 300
accaggcccg cccacaggtag taccgcccc cctgcccatg gtgtcacatc tgcctccagg 360
accaggcctg cattgggtag tacagcaccg ccagtcacaa agttactag tgcctcaggc 420
tcgtctagcg gctcagcttc tactctgggt cacacggcca cctctgcgcg cgcgaccaca 480
acccacagca gcaagagcac tccattctca attccagct gataa 525

<210> 3
 <211> 654
 <212> DNA
 <213> human

<400> 3
 atggggtcag ctctactctt ggtgcacaac ggcacctctg ccagggtctac cacaacccca 60
 gcagagcaaga gcactccatt ctcaattccc agccaccatt ctgatactcc taccacccct 120
 gccagccata gccacnagac tgatgccagt agcaactcacc atagacacgtt auctctctc 180
 aactctccca atcacagcac ttctccccag ttgtctactg gggctctctt ctttttctg 240
 tcttttcaca ttccaaccc tccgtttta tctctctctg aagatccag caccgaactac 300
 taccagagc tgcagagaga cattctgaa atgtttttgc agatttataa acaagggggg 360
 ttctgggccc tctccaatat taagttcagg ccaggatctg ttggtggtaca attgaactgt 420
 gcccttcagag aaggtaccat caatgtccac gacgtggaga cacagttcaa tcagtataaa 480
 accggaagcag cctctcgata taacctgacg atctcagacg tcagcgttag tgatgtgcca 540
 ttctctttct ctgccagtc tggggctggg gtgccaggct ggggcacatgc gctgctggg 600
 ctggtctctg ttctggttgc gctggccatt gtctatctca ttgctctgtg ataa 654

<210> 4
 <211> 285
 <212> DNA
 <213> human

<400> 4
 atgctgggtgc tgggtctgtgt tctggttgcg ctggccattg totatctcat tgccttggt 60
 gtctgtcagt gcgcgcgaaa gaactacggg cagctggaca tctttccaga ccgggatacc 120
 taaccacctca tgagcagta cccacacctac cacaccatg ggcgtatgt gccccctagc 180
 agtaccgatc gtacccccta tgagaaggtt totgcaggtat atggtggcag cagcctctct 240
 tacacaaacc cagcagtggc agccaattct gcccaactgt gataa 285

<210> 5
 <211> 1371
 <212> DNA
 <213> human

<400> 5
 atgacaggtt ctggtcatgc aagctctacc ccagggtggag aaaaggagac ttccgtacc 60
 cagagaagtt cagtgccag ctctactgag aagaatgctg tgagtatgac cagcagcgt 120
 ctctccagcc acagccccc ttcaggctcc tccaccactc agggacagga tgtcaactctg 180
 gccccggcca cggaaaccagc ttccaggttca gctgccacct ggggacagga tgtcaactctg 240
 gctccagtoa ccaggccagc cctgggctcc aaccacccgc cagccacga tgtcaactctc 300
 gccccggaca acaagccagc ccgggaagt accgctccac cagcacacgg tgttaactctg 360
 gctccggata ccaggccggc ccaggttagt accgccctc ctgccatgg tgtcaactct 420
 gccccggaca acaggctcgt attgggtagt acagcaccgc cagtacacaa cgttactagt 480
 gccctaggct ctgctagcgg ctaagctctc actctgggtg acaacggcac ctctgcgcgc 540

```

gagcaccacaa cccacgcgag caagagcact ccattctcaa ttcccagcaa ccactctgat 600
actcctacca cccttgccag ccatagcacc aagactgatg ccagtagcac tcccatatgc 660
acggtaacct cctcacctcc ctccaatcac agcaattctc cccagttgtc tactgggggc 720
ctttcttttt tctgttttt tcaatttca aacctccagt ttaattcttc tctgggaagt 780
cccgaccagg actactacca agagctgcag agagacattt ctgaatgttt tttgcagatt 840
tataaacaag ggggttttct gggcctctcc aatattaagt tcaggccagg atctgtgggtg 900
gtacaattga ctctggcctt ccgagaaggt accatcaatg tccacgagt ggagacacag 960
tccaatcagt ataaaaagg agcagcctct ccatataacc tgacgatctc agacgtcagc 1020
gtgagtgatg tgcatttcc ttctctgcc cagctcgggg ctgggggtgc agcgtggggc 1080
atcgcgctgc tgggtgctgt ctgtgttctg gttgcgctgg ccattgtcta tctcattgcc 1140
ttggctgtct gtcagtggc cggaaagAAC taaggcgagc tggacatctt tccagccggc 1200
gatacctacc atcctatgag cgagtacccc acctaccaca cccatggggc ctatgtgtcc 1260
cctagcagta cgcagtgtag cccctatgag aaggtttctg caggtaatgg tggcagcagc 1320
ctctcttaca caaacccagc agtggcagcc acctctgcca acttgtgata a 1371

```

<210> 6
<211> 369
<212> DNA
<213> human

```

<400> 6
atgcagatct tctgtaagac cctgactggt aagaccatca ctctcgaagt ggagcccgagt 60
gacaccattg agaattgtcaa ggcnaagatc caagacaagg aaggcatccc tcttgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactcttcc tgactacaa 180
atccagaagg agtccacctt gcacctgggt ctccgtctca gaggtgggag gcacggtagt 240
ggtgcactgg tgttgccagt ctgcctgggt aaaagaaaaa cccacctggc gcccaatacg 300
caaacgcgct ctcccgcgog gttggccgat tcattaatgc agctggcagc acaggtttcc 360
cgaggatcc 369

```

<210> 7
<211> 579
<212> DNA
<213> human

```

<400> 7
atgcagatct tctgtaagac cctgactggt aagaccatca ctctcgaagt ggagcccgagt 60
gacaccattg agaattgtcaa ggcnaagatc caagacaagg aaggcatccc tcttgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactcttcc tgactacaa 180
atccagaagg agtccacctt gcacctgggt ctccgtctca gaggtgggag gcacggtagt 240
ggtgcactgg tgttgccagt ctgcctgggt aaaagaaaaa cccacctggc gcccaatacg 300
caaacgcgct ctcccgcgog gttggccgat tcattaatgc agctggcagc acaggtttcc 360
cgaggatcca caggttctgg tcatgcaagc tctaccagc gtggagaaaa ggagactctg 420
gctaccaga gaagttcagt gccacgctct actgagaaga atgctgtgag tatgaccagc 480
agcgtactct ccagccacag ccccggttca ggcctctcca ccaactaggg acaggatgtc 540
actctggccc cggccacgga accagcttca ggttgataa 579

```

<210> 8
 <211> 891
 <212> DNA
 <213> human

<400> 8
 atgcagatct tctgtaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
 gacaccattg agaatgtcaa ggcaagatc caagacaagg aaggcatccc tctgaccag 120
 cagaggctca tctttgcagg caagcagctg gaagatggcc gcactcttcc tgactacaac 180
 atccagaagg agtccacct gcacctggtg ctccgtctca gaggtgggag gacaggtagt 240
 ggtgcatggc tgttgcccggt ctccgtgggtg aaaagaaaaa ccacctggc gcccaatacg 300
 caaacgcgct ctccccgcgc gttggcgat tcattaatgc agctggcag acaggtttcc 360
 cgaggatccg tgcccagctc tactgagaag aatgctgtga gtatgaccag cagcgtactc 420
 tccagocaca gccccgggttc aggtccctcc accactcagg gacaggatgt cactctgtcc 480
 cgggocaggc aaccagcttc aggttcagct gccacctggg gacaggatgt cactcgtg 540
 ccagtaccac ggccagccct gggctccacc accccgccag cccacgatgt cactcagc 600
 ccggacaaca agccagcccc ggggaagtact gctccaccag cacacgggtg tacctcggtc 660
 ccggtatcca ggccggcccc aggtagtacc gccctctctg cccatgggtg cacatctgcc 720
 ccggacaaca ggccctgaatt gggtagtaca gcaccgccag tacacaacgt tactagtccc 780
 tcaggctctg ctacgggctc agcttctact ctgggtcaca accgcaacct tgcgcgcg 840
 accacaacct cagcgagcaa gacgactcca ttctcaattc ccagctgata a 891

<210> 9
 <211> 1020
 <212> DNA
 <213> human

<400> 9
 atgcagatct tctgtaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
 gacaccattg agaatgtcaa ggcaagatc caagacaagg aaggcatccc tctgaccag 120
 cagaggctca tctttgcagg caagcagctg gaagatggcc gcactcttcc tgactacaac 180
 atccagaagg agtccacct gcacctggtg ctccgtctca gaggtgggag gacaggtagt 240
 ggtgcatggc tgttgcccggt ctccgtgggtg aaaagaaaaa ccacctggc gcccaatacg 300
 caaacgcgct ctccccgcgc gttggcgat tcattaatgc agctggcag acaggtttcc 360
 cgaggatccg gctcagcttc tactctgggtg cacacggcag cctctgccag ggctaccaca 420
 accccagcca gcaagagcac tcattctca attccagcc accactctga tactctacc 480
 acccttgcca gccatagcac caagaactgat gccagtatga ctaccatag caaggtacct 540
 ctctcactcc ctcccaatca cagcacttct cccagttgt ctactgggtt ctcttctttt 600
 tctctgtctt ttcaatttc aaacctccag ttttaattct ctctggaga tccagcacc 660
 gactactacc aagagctgca gagagacatt tctgaaatgt ttttgcatg ttatanaaca 720
 ggggggtttt tgggctcttc caatattaag ttccagccag gatctgtgtt ggtacaattg 780
 actctggcct tccgagaagg taccatcaat gtccacgagc tggagacaca gttcaatcag 840
 tataaaacgg aagcagcttc tcgataaac ctgacgatct cagacgtcag cgtgagtgat 900
 gtgccatttc cttctctgce ccagctctggg gctgggggtc caggctgggg catcgcgctg 960
 ctggtgtgtg tctgtgttct ggttgcgctg gccattgtct atctcattgc cttgtgataa 1020

<210> 10
 <211> 651
 <212> DNA
 <213> human

<400> 10
 atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
 gacacatttg agaattgtcaa ggcaaaagatc caagacaagg aaggcatccc tcttgaccag 120
 cagaggctca tctttgcagg caagcagctg gaagatggcc gcactcttcc tgactacaac 180
 atccagaaag agtccaccct gcacctgggt ctccgtctca gaggtggagg gcacggtagt 240
 ggtgcattgg tgttgcccggt ctccgtgggt aaaagaaaaa ccaccctggc gcccaatcgc 300
 caaacgcgct ctccccgcgc gttggcgcat tcattaatgc agctggcagc acaggtttcc 360
 cgaggatccc tgggtcgtgt ctgtgttctg gttgcgtgg ccattgtcta tctcattgac 420
 ttggtcgtct tgcagtgccg ccgaagaagc tacgggcagc tggacatctt tccagccggc 480
 gatactacc atctctatgag cgagtacccc acctaccaca cccatggggc ctatgtgcgc 540
 cctagcagta ccgatcgtag cccctatgag aaggtttctg caggtaatgg tggcagcagc 600
 ctctcttaca caaacccagc agtggcagcc actctcgcca acttggtata a 651

<210> 11
 <211> 1737
 <212> DNA
 <213> human

<400> 11
 atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
 gacacatttg agaattgtcaa ggcaaaagatc caagacaagg aaggcatccc tcttgaccag 120
 cagaggctca tctttgcagg caagcagctg gaagatggcc gcactcttcc tgactacaac 180
 atccagaaag agtccaccct gcacctgggt ctccgtctca gaggtggagg gcacggtagt 240
 ggtgcattgg tgttgcccggt ctccgtgggt aaaagaaaaa ccaccctggc gcccaatcgc 300
 caaacgcgct ctccccgcgc gttggcgcat tcattaatgc agctggcagc acaggtttcc 360
 cgaggatccc caggttctgg tcatgcaagc tatccccag gtggagaaaa ggagactctg 420
 gctaccacaga gaagttcagt gccacgtctc actgagaaga atgctgtgag tatgaccagc 480
 agcgtactct ccagccacag ccccggttca ggctctctca ccactcaggg acaggtatgc 540
 actctggtccc cggccacgga accagcttca ggttcagctg ccacctgggg acaggtatgc 600
 acctcgttcc oagtoaccag gccagccctg ggtctcacca ccccgccagc ccacgatgtc 660
 acctcagccc cggacaacaa gccagccccc ggaagtaccg ctccaccagc acacggtgtt 720
 acctcgggtc cggataccag gccggcccca ggtagtaccg cccctctctg ccatgggtgc 780
 acatctgccc cggacaacag gccctcattg ggtagtacag caccgcaagt acacaaogtt 840
 actagtgcct caggctctgc tagcggctca gcttctatct tgggtgcacaa cggcacctct 900
 gcgcgcgaga ccacaacccc agcagagcaag agcactccat tctcaattcc cagccaccac 960
 tctgatactc ctacacacct tggcagccat agcaccagga ctgatgccag tagcatcac 1020
 catagccctg tacctctctc caccctctcc aatcacagca ctctctccca gttgtctact 1080
 ggggtctctt tctttttctc gtcttttcaac atttcaaac tocagtttaa ttcctctctc 1140
 gaagattccca gcaccgacta ctaccaagag ctgcagagag acattttctg aatggttttg 1200
 cagatttata aacaaggggg ttttctgggc ctcccaata ttaagttcag gccaggtatc 1260
 gtggtgtgac aattgactct ggccttccga gaaggtacca tcaatgtcca cgaogtggag 1320

acacagttca atcagtataa aacggaagca gcctctcgat ataactgac gatctcagac 1380
 gtcagcgtga gtgatgtgcc atttcccttc tctgccagat ctggggctgg ggtgccagcg 1440
 tggggcatcg cgctgctggg gctggctgtg ttctgtgttg cgctggccat tgcctatctc 1500
 attgccttgg ctgtctgtca gtgcgcgca aagaactacg ggcagctgga catctttcca 1560
 gcccgggata cctaccatcc tatgagcgag taacccacct accacaccca tgggcgctat 1620
 gtgcgcccta cgagtacoga tctagacccc tatgagaagg ttctgcagg ttctggcgcc 1680
 agcagctctt cttacacaaa ccagcagtg gcagccaact ctgcccaact gtgataa 1737

<210> 12

<211> 4905

<212> DNA

<213> human

<400> 12

ccagggaagct cctctgtgtc ctcataaacc ctaacotcct ctacttgaga ggacattcca 60
 atcataggct gcccatcacac cotctgtgtc ctctctgttaa ttaggctcaat taacaaaagg 120
 gaaattgggt aggggttttt caacagacgc tttctaaagg taattttaaa atatctggga 180
 agtcccttcc actgctgtgt tccagaagtg ttggtaaaca gcccaacaaat gtcaacagca 240
 gaacataca agctgtcagc ttgacacaag ggcccaacac cctgctcctc aagaagcact 300
 gtggttgtgt gtgtagttaat gtgcaaaa caaggcgcact tttcccccac tgtgtagggt 360
 ccaaaatatt tagtgttttc atttttactt ggatcaggaa cccagcaact cactggataa 420
 gcttatctct tatccaaaac agccttctgtg tcaagtgttca tctgctgact gtcaactgta 480
 gcatttttgg gggttacagt ttgagcagga tatttggctc tgtagtttgc taacacaccc 540
 tgcagctcca aagggtcccc accaacagca aaaaatgaa aatttgaccc ttgnaatgggt 600
 ttccagcagc cattttcatg agttttttgt gtccctgaat gcaaghttaa catagcagtt 660
 accccaataa cctcagtttt aacagtaaca gcttcccaca tcaaaatatt tccacaggtt 720
 aagtctcat taaattagg caaaggaatt cttgaagaag aaagggcctc gtgatacgcc 780
 tatttttata ggttaatgto atgataataa tgggttttta gacgtcaagt ggcaattttc 840
 ggggaaatgt gcgcggaacc cctatttgtt tattttttta aatacattca aatattgtac 900
 cgcctcagag acaataaccc tgataaatgc ttcaataata ttgaaaagg aagagtatga 960
 gtattcaaca tttcctgttc gcccttattc ccttttttgc ggcattttgc ctctcgtttt 1020
 tttgctcacc agaaaacgtg gtgaagtaa aagatgtgta agatcagttg ggtgcacgag 1080
 tgggtttcat cgaactggat ctcaacagcg gtaagatctt tgagagtttt cgcgccgaag 1140
 aacgttttcc aatgatgagc acttttaaa ttctgtatg tggcgcggta ttatcccggt 1200
 ttgacgcggg gcaagagcaa ctccgttgcc gcatacacta tttccagaat gacttggttg 1260
 agtatccacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 1320
 gtgctgccat aacatagat gataaacctg cggccaactt acttctgaca acgatcggag 1380
 gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaaat cgccttgato 1440
 tctgggaacc agagctgaat gaagccatcc caaacgacga cggtgacacc acgatgcctg 1500
 cagcaatggc aacaacgttg cgaataactat taactggcga actacttaet etagcttccc 1560
 gccaacaatt aatgactcgg atggaggcgg ataaagtgc aggaaccact ctgcctcggg 1620
 ccctccggc tggctgtggtt attgtcgata aactctgagc cggctgagct gggtctcggg 1680
 gtatcatctg agcaactggg ccagatggta agccctcccg tatcgtagtt atctacagga 1740
 cggggagctca ggcaactatg gatgaacgaa atagacagat cgttgagata ggtgcctcac 1800
 tgatgaagca ttggtaactg tcaagccaaag ttactcata tatactttag attgatttaa 1860
 aactctattt ttaatttaa aggatctagg tgaagatcct ttttgataat ctcatgacca 1920
 aaatccctta acgtgagttt tctgtccact gagcgtcaga ccccgtagaa aagatcaagg 1980

gatcttcttg agatectttt tttctgcgcg taatctgctg cttgcnaaca aaaaaaccac 2040
 cgctaacacg ggtgggttgt ttgcgggacg aagaagctac aactcttttt ccgaaggtaa 2100
 ctggcttccg cagagcgccg atacaaaata ctgtctctct agtgtagcgg tagttaggcc 2160
 accacttcaa gaactctgtg gcacccgcct cataactcgc tctgtcaatc ctgtttaccg 2220
 tggctctgcg cagtggggat aagtcgtgtc ttacgggggt ggaactcaaa cgaatgttac 2280
 cggataaagg ccagcgggtg ggcgtgaacg ggggttctgt cacacagccc agcttggagc 2340
 gaacgaceta cacogaactg agataoctac agcgtgagct atgagaaago gccacogctc 2400
 ccgaaggagg aaaggcgacg aggtatccgg taagcgggcag ggtcggaaca gggagagcca 2460
 cgaggggagc tcacggggga aacgcctggg atctttatag tctctcggg ttctgcacac 2520
 tctgaattga gctcgtatt ttgtgatgct cgtcaggggg cgggagccta tggaaaacgg 2580
 ccagcaacgc ggccttttta cgggttcctgg ccttttctgt gcccttttgc cacatgttct 2640
 ttctcgtggt atccctgat tctgtggata accgtattac gccttttgag tgaagctgata 2700
 cgcctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa cgggaagagc 2760
 gctgtagcgg gtatttttct cttacgcctc tctgcgggtat ttacacccgc atatggtgca 2820
 ctctcagtcac aatctgctct gatgcgcgat agttaagcca gtatacaate aatattggcg 2880
 attagccata ttattcattg gttatatagc ataactcaat attgctattt ggcatttcca 2940
 taactgtgat ccatatcata atatgtacat ttatatgtgc taatgtccaa cattaccgco 3000
 atgttgacat tgattattga ctagtattta atagtaacca attacggggg cattagtcca 3060
 tagcccatat atggagttcc cgtgtacata acttaacgga aatggccccc ctggctgaac 3120
 gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacggcaat 3180
 agggacttct cattgacgtc aatgggtgga gtatttacgg taactggccc acttggcagt 3240
 acatcaagtg tatcatatgc caagtacgcc cctattgac gtcaatgaag gtaaatggcc 3300
 cgcctgggat tatccccagt acatgacctt atgggaacttt cctacttggc agtacatcta 3360
 cgtattagtc atcgtatta ccatgggtgat gcggttttgg cagtacatca atggcgctgg 3420
 atagcggttt gactcacggg gatttccaaq tctccacccc attgacgtca atgggagttt 3480
 gttttggcac caaaatocac ggggaacttcc aaaaatgctg acaacactcg ccccatgtac 3540
 gccaaatggg ggtaggcggtg tacgggtggga ggtctatata agcagagctc gttttagtga 3600
 ccgtcagatc gccctggagac gccatccacg ctgttttgac ctccatagaa gacacgggga 3660
 ccgatccagc ctcocggccc ggggaacggtg cattgggaacg cggattcccc gtgcnaagca 3720
 agcttctcta gaacccggga gagctctcga gaacttcagg gtgagtttgg ggaacottga 3780
 ttgttctctt tttttcgtca ttgtanaatt catgttatat ggagggggca aagttttcag 3840
 ggtgttgttt agaattgggaa gatgtccctt gttacacat ggacccctcat gataattttg 3900
 ttcttttctc ttctactctt gttgacaacc attgtctctt cttattttct ttctatttct 3960
 tgtacacttt ttgttaaaat ttagcttgca ttgtgaacga atttttaaat tcaacttttg 4020
 ttatttgtca gattgtgaat actttctcta atcaactttt ttctaaagga accagggtat 4080
 attatattgt acttcagcac agtttttagag aacaattgtt ataattaaat gataaggtag 4140
 aatatctctg catataaatt ctggctggcg tggaaaatbt ctatttgga tgaacaacta 4200
 catctcgtgc atcatctcgc cttctctctt atggttacaa tgatatacac tgtttgagat 4260
 gaggataaaa tactctgagt coaaacoggg cccctctgct aaccatgttc atgccttctt 4320
 cttttctcta cagctcctgg goaacgtgct ggttgttgtg ctgtctcatc attttggcca 4380
 agaattcact cctcaggtgc aggtcgccca tcagaaggtg gtggctgggt tggccaatgc 4440
 cctggctcac aaataccact gagatctttt tccctctgcc aaaaattatg gggacatcat 4500
 gaagccocctg gagcatctga cttctggcta ataaaggaaa ttatttttca ttgcaatagt 4560
 gtgttggaaat tttttgtgct tctcactcgg aaggacatat gggaggcgca atcattttaa 4620
 acatcagaat gagtatttgg ttttagagtt ggcaacatat gccatatgct ggtcgcaatg 4680
 aacaaagggt gctataaaga ggtcatcagt atatgaaca gccocctgct gctccattct 4740
 tattccatag aaaagccttg acttgaggtt agattttttt tatattttgt ttgtggttat 4800
 tttttctctt aacatcccta aaattttctt tacatgtttt actagcgaag tttttctctc 4860

WO 00/25827

PCT/EP99/07874

tctcttgact actccagtc atagctgtcc ctctctctg gatcc

4905

<210> 13

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 13

gatcggatcc acaggttctg gtcattgcaag c

31

<210> 14

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 14

gatctctaga aaatttatca acctgaagct gggtccgtgg c

41

<210> 15

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 15

gatcggatcc gtgccagct ctactgagaa gaatgc

36

<210> 16

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 16

gatctctaga aagcttatca gctgggaatt gagaatggag tgctcttgc

49

<210> 17

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 17

gacggatcc ggctcagctt ctactctggt gcacaacggc

40

<210> 18

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 18

gatctctaga aagcttatca caaggcaatg agatngacaa tggcc

45

<210> 19

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 19

gacggatcc ctggtgctgg tctgtgtctt ggttgcc

38

<210> 20

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 20

gatctctaga aagcttatca caagttggca gaagtggctg c

41

<210> 21

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 21

gatctctaga atgacaggtt ctggtcatgc aagc

34

<210> 22

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 22

gatctctaga atgggtgccca gctctactga gaagaatgc

39

<210> 23

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 23

gatctctaga atgggtctcag cttctactct ggtgcacac ggc

43

<210> 24

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 24

gactctctaga atgctgggtgc tggctctgtgt tctggttgcg c

41

<210> 25

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 25

ggcgggtggag cccggggctg gcttgt

26

<210> 26

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 26

aacctgaagc tggttccgtg gc

22

<210> 27

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic

oligonucleotide

<400> 27

gtgccagct ctactgagaa gaatgc

26

<210> 28

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 28

gctgggaatt gagaatggag tgctcttgc

29

<210> 29

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 29

ggctcagctt ctactctggt gcacaacggc

30

<210> 30

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 30

caaggcaatg agatogacaa tggcc

25

<210> 31

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 31

ctggtgtctgg tctgtgttct ggttgcg

27

<210> 32

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 32

gatctctaga atgcagatct tctgaagac cctgactggt

40

<210> 33

<211> 68

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 33

tcaccagcga gacgggcaac agccatgcac cactaccgtg cctccacct ctgagacgga
gcaccagg

68

<210> 34

<211> 66

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 34

ctcccgctcc agaggtggga ggcacggtag tggtagcatgg ctgttgcctg tctcgtggt

60

WO 00/25827

PCT/EP99/07874

gaaaag

66

<210> 35

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
oligonucleotide

<400> 35

gacggatcc tcggaaacc tgtcgtgcc gctgc

35